

METHOD AND DEVICE FOR SOUND CODE CONVERSION**Patent number:** JP2003076394**Publication date:** 2003-03-14**Inventor:** TSUCHINAGA YOSHITERU; OTA TAKASHI; SUZUKI MASANAO**Applicant:** FUJITSU LTD**Classification:**

- international: G10L19/00; G10L19/04; H03M7/36; H04J3/00

- european:

Application number: JP20010263031 20010831**Priority number(s):****Also published as:**

EP1288913 (A2)

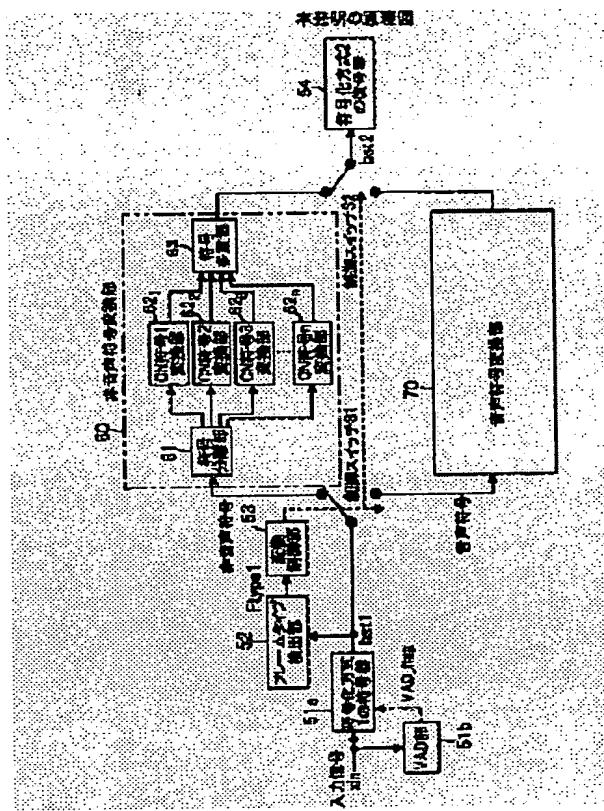
US2003065508 (/)

EP1288913 (A3)

Abstract of JP2003076394

PROBLEM TO BE SOLVED: To convert a non-sound code (CN code) encoded by an encoding method on the transmission side to a non-sound code conforming to an encoding method on the reception side without decoding to a CN signal.

SOLUTION: A first non-sound code obtained by encoding a non-sound signal included in an input signal by the non-sound compressing function of a first sound encoding system is converted to a second non-sound code of a second sound encoding system without being temporarily decoded to a non-sound signal. For example, the first non-sound code is separated into a plurality of first element codes by a code separation part 61, and the first element codes are converted to a plurality of second element codes constituting the second non-sound code by CN code conversion parts 621 to 62n, and the second element codes obtained by this conversion are multiplexed to output the second non-sound code.

Report a data error

Data supplied from the esp@cenet database - Patent Abstracts of Japan

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2003-76394

(P2003-76394A)

(43)公開日 平成15年3月14日 (2003.3.14)

(51) Int.Cl.⁷

識別記号

F I

テ-マコ-ト(参考)

G 10 L 19/00

H 03 M 7/36

5 D 0 4 5

19/04

H 04 J 3/00

S 5 J 0 6 4

H 03 M 7/36

G 10 L 9/18

A 5 K 0 2 8

H 04 J 3/00

9/14

J

審査請求 未請求 請求項の数10 O L (全 25 頁)

(21)出願番号 特願2001-263031(P2001-263031)

(71)出願人 000005223

富士通株式会社

(22)出願日 平成13年8月31日 (2001.8.31)

神奈川県川崎市中原区上小田中4丁目1番

1号

(72)発明者 土永 義照

福岡県福岡市博多区博多駅前三丁目22番8

号 富士通九州ディジタル・テクノロジ株
式会社内

(74)代理人 100084711

弁理士 斎藤 千幹

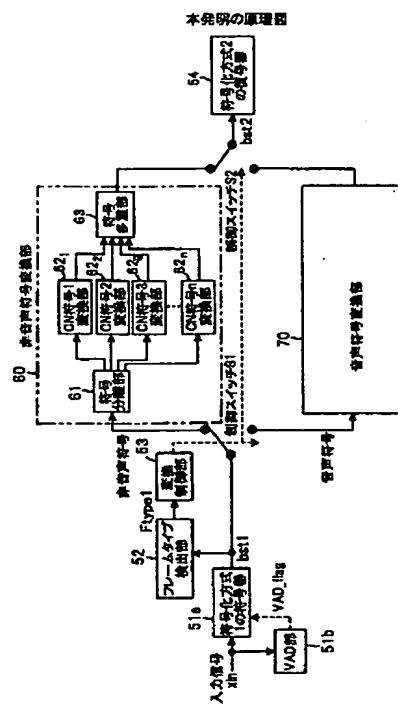
最終頁に続く

(54)【発明の名称】 音声符号変換方法及び装置

(57)【要約】

【課題】 送信側の符号化方法で符号化した非音声符号(CN符号)をCN信号に復号しなくても受信側の符号化方法に応じた非音声符号に変換する。

【解決手段】 入力信号に含まれる非音声信号を第1の音声符号化方式の非音声圧縮機能により符号化して得られた第1の非音声符号を一旦非音声信号に復号することなく第2の音声符号化方式の第2の非音声符号に変換する。例えば、符号分離部61で第1の非音声符号を第1の複数の要素符号に分離し、CN符号変換部62₁～62nで第1の各要素符号を第2の非音声符号を構成する第2の複数の要素符号に変換し、この変換により得られた第2の各要素符号を多重して第2の非音声符号を出力する。



【特許請求の範囲】

【請求項1】 入力信号を第1の音声符号化方式で符号化して得られる第1の音声符号を、第2の音声符号化方式の第2の音声符号に変換する音声符号変換方法において、
入力信号に含まれる非音声信号を第1の音声符号化方式の非音声圧縮機能により符号化して得られた第1の非音声符号を一旦非音声信号に復号することなく第2の音声符号化方式の第2の非音声符号に変換する、
ことを特徴とする音声符号変換方法。

【請求項2】 入力信号を第1の音声符号化方式で符号化して得られる第1の音声符号を、第2の音声符号化方式の第2の音声符号に変換する音声符号変換方法において、
入力信号に含まれる非音声信号を第1の音声符号化方式の非音声圧縮機能により符号化して得られた第1の非音声符号を第1の複数の要素符号に分離し、第1の複数の要素符号を前記第2の非音声符号を構成する第2の複数の要素符号に変換し、
前記変換により得られた第2の複数の要素符号を多重化して第2の非音声符号を出力する、
ことを特徴とする音声符号変換方法。

【請求項3】 前記変換ステップにおいて、前記第1の複数の要素符号を第1の音声符号化方式と同じ量子化テーブルを持つ逆量子化器で逆量子化し、
逆量子化により得られた複数の要素符号の逆量子化値を第2の音声符号化方式と同じ量子化テーブルを持つ量子化器で量子化して第2の複数の要素符号に変換する、
ことを特徴とする請求項2記載の音声符号変換方法。

【請求項4】 入力信号の一定サンプル数をフレームとし、フレーム単位で音声区間における音声信号を第1の音声符号化方式で符号化して得られる第1の音声符号と、
非音声区間における非音声信号を第1の非音声符号化方式で符号化して得られる第1の非音声符号を混在して送信側より伝送し、これら第1の音声符号と第1の非音声符号をそれぞれ、第2の音声符号化方式による第2の音声符号と第2の非音声符号化方式による第2の非音声符号とにそれぞれ変換し、変換により得られた第2の音声符号と第2の非音声符号を混在して受信側に伝送する音声通信システムにおける音声符号変換方法において、
非音声区間では所定のフレームにおいてのみ非音声符号を伝送し、それ以外のフレームでは非音声符号を伝送せず、

前記フレーム単位の符号情報に、音声フレーム、非音声フレーム、符号を伝送しない非伝送フレームの別を示すフレームタイプ情報を付加し、フレームタイプ情報に基いてどのフレームの符号であるか識別し、非音声フレーム、非伝送フレームの場合には、第1、第2の非音声符号化方式におけるフレーム長の差、および非音声符号の伝送制御の相違を考慮して第1の非音声符号を第2の非音声符号に

変換する、

ことを特徴とする音声符号変換方法。

【請求項5】 (1)第1の非音声符号化方式が、非音声区間における所定フレーム数毎に平均した非音声符号を伝送すると共に、その他のフレームでは非音声符号を伝送しない方式であり、(2)第2の非音声符号化方式が、非音声区間における非音声信号の変化の度合が大きいフレームにおいてのみ非音声符号を伝送し、その他のフレームでは非音声符号を伝送せず、しかも、連続して非音声符号を伝送しない方式であり、更に、(3)第1の非音声符号化方式のフレーム長が、第2の非音声符号化方式のフレーム長の2倍であるとき、
第1の非音声符号化方式における非伝送フレームの符号情報を第2の非音声符号化方式における2つの非伝送フレームの符号情報に変換し、第1の非音声符号化方式における非音声フレームの符号情報を、第2の非音声符号化方式における非音声フレームの符号情報と非伝送フレームの符号情報との2つに変換する、ことを特徴とする請求項6記載の音声符号変換方法。

【請求項6】 音声区間から非音声区間に変化するとき、前記第1の非音声符号化方式が、変化点のフレームを含めて連続nフレームは音声フレームとみなして音声符号を伝送し、次のフレームは非音声符号を含まない最初の非音声フレームとしてフレームタイプ情報を伝送する場合、第1の非音声符号化方式における前記最初の非音声フレームが検出された時、第1の音声符号化方式における直前n個の音声フレームの音声符号を逆量子化して得られる逆量子化値を平均化し、平均値を量子化して前記第2の非音声符号化方式の非音声フレームにおける非音声符号を求める、

ことを特徴とする請求項7記載の音声符号変換方法。

【請求項7】 (1)第1の非音声符号化方式が、非音声区間における非音声信号の変化の度合が大きいフレームにおいてのみ非音声符号を伝送し、その他のフレームでは非音声符号を伝送せず、また、連続して非音声符号を伝送しない方式であり、(2)第2の非音声符号化方式が、非音声区間における所定フレーム数N毎に平均した非音声符号を伝送すると共に、その他のフレームでは非音声符号を伝送しない方式であり、更に、(3)第1の非音声符号化方式のフレーム長が、第2の非音声符号化方式のフレーム長の半分であるとき、

第1の非音声符号化方式の連続する2×Nフレームにおける各非音声符号の逆量子化値を平均し、平均値を逆量子化して第2の非音声符号化方式におけるNフレーム毎のフレームの非音声符号とし、Nフレーム毎以外のフレームについては、第1の非音声符号化方式の連続する2つのフレームの符号情報をフレームタイプに関係なく第2の非音声符号化方式の1つの非伝送フレームの符号情報に変換する、ことを特徴とする請求項6記載の音声符号変換方法。

【請求項8】 音声区間から非音声区間に変化するとき、前記第2の非音声符号化方式が、変化点のフレームを含めて連続nフレームは音声フレームとみなして音声符号を伝送し、次のフレームは非音声符号を含まない最初の非音声フレームとしてフレームタイプ情報を伝送する場合、第1の非音声フレームの非音声符号を逆量子化して複数の要素符号の逆量子化値を発生し、同時に、予め定めた、あるいはランダムな別の要素符号の逆量子化値を発生し、連続する2フレームの各要素符号の逆量子化値を第2音声符号化方式の量子化テーブルを用いてそれぞれ量子化して第2音声符号化方式の1フレーム分の音声符号に変換し、nフレーム分の第2音声符号化方式の音声符号を出力した後、非音声符号を含まない前記最初の非音声フレームのフレームタイプ情報を送出する、ことを特徴とする請求項9記載の音声符号変換方法。

【請求項9】 入力信号を第1の音声符号化方式で符号化して得られる第1の音声符号を、第2の音声符号化方式の第2の音声符号に変換する音声符号変換装置において、

入力信号に含まれる非音声信号を第1の音声符号化方式の非音声圧縮機能により符号化して得られた第1の非音声符号を第1の複数の要素符号に分離する符号分離部、第1の複数の要素符号を、前記第2の非音声符号を構成する第2の複数の要素符号に変換する要素符号変換部、前記変換により得られた第2の各要素符号を多重化して第2の非音声符号を出力する符号多重部、

を備えたことを特徴とする音声符号変換装置。

【請求項10】 入力信号の一定サンプル数をフレームとし、フレーム単位で音声区間における音声信号を第1の音声符号化方式で符号化して得られる第1の音声符号と、非音声区間における非音声信号を第1の非音声符号化方式で符号化して得られる第1の非音声符号を混在して送信側より伝送し、これら第1の音声符号と第1の非音声符号をそれぞれ、第2の音声符号化方式による第2の音声符号と第2の非音声符号化方式による第2の非音声符号とにそれぞれ変換し、変換により得られた第2の音声符号と第2の非音声符号を受信側に伝送する音声通信システムにおける音声符号変換装置において、

符号情報に付加されているフレームタイプ情報に基いて、音声フレーム、非音声フレーム、非音声区間において非音声符号を伝送しない非伝送フレームの別を識別するフレームタイプ識別部、非音声フレームにおける第1の非音声符号を、第1の非音声符号化方式と同じ量子化テーブルに基いて逆量子化し、得られた逆量子化値を第2の非音声符号化方式と同じ量子化テーブルに基いて量子化して第2の非音声符号に変換する非音声符号変換部、第1、第2の非音声符号化方式におけるフレーム長の差、および非音声符号の伝送制御の相違を考慮して前記非音声符号変換部を制御する変換制御部、を有することを特徴とする音声符号変換装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は音声符号変換方法及び装置に係わり、特に、インターネットなどのネットワークで用いられる音声符号化装置や自動車・携帯電話システム等で用いられる音声符号化装置によって符号化された音声符号を別の符号化方式の音声符号に変換する音声符号変換方法及び装置に関する。

【0002】

【従来の技術】近年、携帯電話加入者が爆発的に増加しており、今後も増加し続けることが予想される。また、インターネットを使った音声通信(Voice over IP:VoIP)は、企業内ネットワークや長距離電話サービスなどの分野で普及してきている。このような音声通信システムでは、通信回線を有効利用するため音声を圧縮する音声符号化技術が用いられるが、システム毎に使用される音声符号化方式が異なる。例えば、次世代の携帯電話システムとして期待されているW-CDMAでは、世界共通の音声符号化方式としてAMR(Adaptive Multi-Rate;適応マルチレート)方式が採用されている。一方、VoIPでは音声符号化方式としてITU-T勧告G.729A方式が広く用いられている。

【0003】今後、インターネットと携帯電話の普及に伴い、インターネットユーザーと携帯電話ユーザによる音声通信の通信量がますます増加すると考えられる。ところが、前述したように携帯電話網とインターネット網では、使用する音声符号化方式が異なるためそのままで通信することができない。このため、従来は一方のネットワークで符号化された音声符号を音声符号変換器により、他方のネットワークで使用されている音声符号方式の音声符号に変換する必要がある。

【0004】音声符号変換図15に従来の典型的な音声符号変換方法の原理図を示す。以下ではこの方法を従来技術1と呼ぶ。図において、ユーザAが端末1に対して入力した音声をユーザBの端末2に伝える場合のみを考える。ここで、ユーザAの持つ端末1は符号化方式1の符号器1aのみを持ち、ユーザBの持つ端末2は符号化方式2の復号器2aのみを持つこととする。

【0005】送信側のユーザAが発した音声は、端末1に組み込まれた符号化方式1の符号器1aへ入力する。符号器1aは入力した音声信号を符号化方式1の音声符号に符号化して伝送路1bに送出する。音声符号変換部3の復号器3aは、伝送路1bを介して音声符号が入力すると、符号化方式1の音声符号から一旦再生音声を復号する。続いて、音声符号変換部3の符号器3bは再生音声信号を符号化方式2の音声符号に変換して伝送路2bに送出する。この符号化方式2の音声符号は伝送路2bを通して端末2に入力する。復号器2aは音声符号が入力すると、符号化方式2の音声符号から再生音声を復号する。これにより、受信側のユーザBは再生音声を聞

くことができる。以上のように一度符号化された音声を復号し、復号された音声を再度符号化する処理をタンドム接続と呼ぶ。

【0006】以上のように従来技術1の構成では、音声符号化方式1で符号化した音声符号を一旦符号化音声に復号し、再度、音声符号化方式2により符号化するタンドム接続を行うため、音声品質の著しい劣化や遅延の増加といった問題があった。このようなタンドム接続の問題点を解決する方法として、音声符号を音声信号に戻すことなく、LSP符号、ピッチラグ符号等のパラメータ符号に分解し、各パラメータ符号を個別に別の音声符号化方式の符号に変換する手法が提案されている（特願2001-75427参照）。図16にその原理図を示す。以下ではこれを従来技術2と呼ぶ。

【0007】端末1に組み込まれた符号化方式1の符号器1aはユーザAが発した音声信号を符号化方式1の音声符号に符号化して伝送路1bに送出する。音声符号変換部4は伝送路1bより入力した符号化方式1の音声符号を符号化方式2の音声符号に変換して伝送路2bに送出し、端末2の復号器2aは、伝送路2bを介して入力する符号化方式2の音声符号から再生音声を復号し、ユーザBはこの再生音声を聞くことができる。

【0008】符号化方式1は、①フレーム毎の線形予測分析により得られる線形予測係数（LPC計数）から求まるLSPパラメータを量子化することにより得られる第1のLSP符号と、②周期性音源信号を出力するための適応符号帳の出力信号を特定する第1のピッチラグ符号と、③雑音性音源信号を出力するための代数符号帳（あるいは雑音符号帳）の出力信号を特定する第1の代数符号（雑音符号）と、④前記適応符号帳の出力信号の振幅を表すピッチゲインと前記代数符号帳の出力信号の振幅を表す代数ゲインとを量子化して得られる第1のゲイン符号とで音声信号を符号化する方式である。又、符号化方式2は、第1の音声符号化方式と異なる量子化方法により量子化して得られる①第2のLSP符号、②第2のピッチラグ符号、③第2の代数符号（雑音符号）、④第2のゲイン符号とで音声信号を符号化する方式である。

【0009】音声符号変換部4は、符号分離部4a、LSP符号変換部4b、ピッチラグ符号変換部4c、代数符号変換部4d、ゲイン符号変換部4e、符号多重化部4fを有している。符号分離部4aは、端末1の符号器1aから伝送路1bを介して入力する符号化方式1の音声符号より、音声信号を再現するために必要な複数の成分の符号、すなわち、①LSP符号、②ピッチラグ符号、③代数符号、④ゲイン符号に分離し、それぞれを各符号変換部4b～4eに入力する。各符号変換部4b～4eは入力された音声符号化方式1によるLSP符号、ピッチラグ符号、代数符号、ゲイン符号をそれぞれ音声符号化方式2によるLSP符号、ピッチラグ符号、代数符号、ゲイン符号に変換し、符号多重化部4fは変換された音声符

号化方式2の各符号を多重化して伝送路2bに送出する。

【0010】図17は各符号変換部4b～4eの構成を明示した音声符号変換部の構成図であり、図16と同一部分には同一符号を付している。符号分離部4aは伝送路より入力端子#1を介して入力する符号化方式1の音声符号より、LSP符号1、ピッチラグ符号1、代数符号1、ゲイン符号1を分離し、それぞれ符号変換部4b～4eに入力する。

【0011】LSP符号変換部4bのLSP逆量子化器4b₁は、符号化方式1のLSP符号1を逆量子化してLSP逆量子化値を出力し、LSP量子化器4b₂は該LSP逆量子化値を符号化方式2のLSP量子化テーブルを用いて量子化してLSP符号2を出力する。ピッチラグ符号変換部4cのピッチラグ逆量子化器4c₁は、符号化方式1のピッチラグ符号1を逆量子化してピッチラグ逆量子化値を出力し、ピッチラグ量子化器4c₂は該ピッチラグ逆量子化値を符号化方式2のピッチラグ量子化テーブルを用いて量子化してピッチラグ符号2を出力する。代数符号変換部4dの代数符号逆量子化器4d₁は、符号化方式1の代数符号1を逆量子化して代数符号逆量子化値を出力し、代数符号量子化器4d₂は該代数符号逆量子化値を符号化方式2の代数符号量子化テーブルを用いて量子化して代数符号2を出力する。ゲイン符号変換部4eのゲイン逆量子化器4e₁は、符号化方式1のゲイン符号1を逆量子化してゲイン逆量子化値を出力し、ゲイン量子化器4e₂は該ゲイン逆量子化値を符号化方式2のゲイン量子化テーブルを用いて量子化してゲイン符号2を出力する。符号多重化部4fは、各量子化器4b₂～4e₂から出力するLSP符号2、ピッチラグ符号2、代数符号2、ゲイン符号2を多重して符号化方式2による音声符号を作成して出力端子#2より伝送路に送出する。

【0012】図15のタンドム接続方式（従来技術1）は、符号化方式1で符号化された音声符号を一旦音声に復号して得られた再生音声を入力とし、再度符号化と復号を行っている。このため、再度の符号化（つまり音声情報圧縮）によって原音に比べて遙かに情報量が少なくなっている再生音声から音声のパラメータ抽出を行うため、それによって得られる音声符号は必ずしも最適なものではなかった。これに対し、図16の従来技術2の音声符号化装置によれば、符号化方式1の音声符号を逆量子化及び量子化の過程を介して符号化方式2の音声符号に変換するため、従来技術1のタンドム接続に比べて格段に劣化の少ない音声符号変換が可能となる。また、音声符号変換のために一度も音声に復号する必要がないので、従来のタンドム接続で問題となっていた遅延も少なくて済むという利点がある。

【0013】非音声圧縮
ところで、実際の音声通信システムは、音声会話に含まれる非音声区間を有効利用してさらに情報の伝送効率を

向上させる非音声圧縮機能を持つのが一般的である。図18に非音声圧縮機能の概念図を示す。人の会話では、音声と音声の間に無音部、背景雑音部などの非音声区間が存在する。このような区間では音声情報を伝送する必要が無く、通信回線をより有効利用できる。これが非音声圧縮の基本的な考え方である。しかし、このままでは受信側で再生された音声と音声の間が全くの無音になり聴覚的に不自然さが生じるため、通常は受信側で聴覚的に違和感のない自然なノイズ(コンフォートノイズ)を発生させる。入力信号に類似したコンフォートノイズを生成するため、送信側よりコンフォートノイズ情報(以下、CN情報と呼ぶ)を伝送する必要があるが、CN情報の情報量は音声に比べ少なく、また非音声区間の性質は緩やかに変化するため常にCN情報を送る必要がない。これにより音声区間に比べ伝送する情報量を大幅に低減できるため、通信回線全体の伝送効率をさらに向上させることができる。このような非音声圧縮機能は、音声区間・非音声区間を検出するVAD部(Voice Activity Detection:音声区間検出)、送信側でCN情報の生成・伝送制御を行うDTX部(Discontinuous Transmission:不連続伝送制御)、受信側でコンフォートノイズを発生させるCNG部(Comfort Noise Generator:コンフォートノイズ発生器)で実現される。

【0014】以下、非音声圧縮機能の動作原理を説明する。図19に原理図を示す。送信側において、一定長のフレーム(例えば、80サンプル/10msec)に分割した入力信号をVAD部5aに入力して音声区間検出を行う。VAD部5aは、音声区間で1、非音声区間で0の判定結果vad_flagを出力する。音声区間(vad_flag=1)の場合、スイッチSW1～SW4をすべて音声側に切り替え、送信側の音声符号器5b及び受信側の音声復号器6aは通常の音声符号化方式(例えば、G.729AやAMR)にしたがって音声信号の符号化、復号化を行う。一方、非音声区間(vad_flag=0)の場合、スイッチSW1～SW4をすべて非音声側に切り替え、送信側の非音声符号器5cはDTX部(図示せず)の制御で非音声信号の符号化処理、すなわち、CN情報の生成・伝送制御を行い、受信側の非音声復号器6bはCNG部(図示せず)の制御で復号化処理、すなわち、コンフォートノイズを発生する。

【0015】次に非音声符号器5c、非音声復号器6bの動作について説明する。図20にそれぞれのブロック図、図21(a), (b)にそれぞれの処理フローを示す。

CN情報生成部7aでは、フレーム毎に入力信号を分析して受信側のCNG部8aでコンフォートノイズを生成するためのCNパラメータを算出する(ステップS101)。CNパラメータとしては一般的に周波数特性の概形情報と振幅情報が用いられる。DTX制御部7bはスイッチ7cを制御して、求めたCN情報を受信側へ伝送する/しないをフレーム毎に制御する(S102)。制御方法としては、信号の性質に応じて適応的に制御する方法や一定間隔で定期的に制御する方法がある。伝送が必要な場合には、CNパラメータをCN量子化部7dへ入力し、CN量子化部7dはCNパラメータを量子化してCN符号を生成し(S103)、回線データとして受信側へ伝送する(S104)。以後、CN情報が伝送されるフレームをSID(Silence Insertion Descriptor)フレームと呼ぶ。その他のフレームでは、非伝送フレームとなり何も伝送しない(S105)。

【0016】受信側のCNG部8aは、伝送してきたCN符号を基にコンフォートノイズを発生する。すなわち、送信側から送られてきたCN符号は、CN逆量子化部8bに入力し、CN逆量子化部8bは該CN符号を逆量子化してCNパラメータにし(S111)、CNG部8aはCNパラメータを用いてコンフォートノイズを生成(S112)する。また、CNパラメータが伝送されて来ない非伝送フレームでは、最後に受信したCNパラメータを用いてコンフォートノイズを生成する(S113)。以上のように、実際の音声通信システムでは、会話の中の非音声区間を判別し、この非音声区間において受信側で聴覚的に自然なノイズを生成するための情報のみを間欠的に伝送し、これにより伝送効率をさらに向上させることが可能である。このような非音声圧縮機能は、先に述べた次世代携帯電話網やVoIP網でも採用されており、システム毎に異なる方式が用いられている。

【0017】次に代表的な符号化方式であるG.729A(VoIP)とAMR(次世代携帯電話)に用いられている非音声圧縮機能について説明する。表1に両方式の諸元を示す。

【表1】

G.729AとAMRの非音声圧縮機能の比較

| | | G.729A | AMR |
|-------------|--------------------|--|-------------------------------------|
| 処理フレーム長 | | 10ms(80sample) | 20ms(160sample) |
| 伝送される CN 情報 | LPC 係数 フレーム信号電力 | LPC 係数 フレーム信号電力 | LPC 係数 フレーム信号電力 |
| CN 情報の生成方法 | LPC 情報 | 過去 6 フレームの平均 LPC 係数または現フレームの LPC 係数 | 過去 8 フレームの平均 LPC 係数(LSP 領域で算出) |
| | フレーム信号電力情報 | 過去 0~3 フレームの平均 対数電力(LSP 残差信号領域) | 過去 8 フレームの平均対数電力(入力信号領域) |
| | CN 符号のビット割り当て | 10bit(LSP 領域で量子化) フレーム信号電力 合計 | 29bit(LSP 領域で量子化) 6bit 35bit |
| DTX 制御方法 | | ・適応制御(非音声信号の変動に応じて不定期に伝送) | ・固定制御(8 フレーム毎に定期的に伝送) ・ハングオーバー制御 |

G.729A、AMRともCN情報としてLPC係数(線形予測計数)とフレーム信号電力が用いられる。LPC係数は入力信号の周波数特性の概形を表わすパラメータであり、フレーム信号電力は入力信号の振幅特性を表わすパラメータである。これらパラメータはフレーム毎に入力信号を分析することによって得られる。以下にG.729AとAMRのCN情報の生成方法を述べる。

【0018】G.729Aでは、LPC情報は現フレームを含む過去6フレームのLPC係数の平均値として求められる。また、SIDフレーム近傍の信号変動を考慮して、求めた平均値または現フレームのLPC係数を最終的にCN情報として用いる。どちらを選択するかは、両LPC係数間のひずみを測定することによって決定される。信号に変動がある(歪が大きい)と判定された場合、現フレームのLPC係数が用いられる。フレーム電力情報は、LPC予測残差信号の対数電力を現フレームを含む過去0~3フレームで平均化した値として求められる。ここでLPC残差信号は、フレーム毎に入力信号をLPC逆フィルタに通すことによって得られる信号である。

【0019】AMRでは、LPC情報は現フレームを含む過去8フレームのLPC係数の平均値として求められる。平均値の算出はLPC係数をLSPパラメータに変換した領域で行われる。ここで、LSPはLPC係数と相互に変換が可能な周波数領域のパラメータである。フレーム信号電力情報は、入力信号の対数電力を過去8フレーム(現フレームを含む)で平均化した値として求められる。以上のようにG.729A、AMRともCN情報としてLPC情報とフレーム信号電力情報を用いるが、その生成(算出)方法は異なる。

【0020】CN情報はCN符号に量子化され復号器へと伝送される。表1にG.729AとAMRのCN符号のビット割り当てを示す。G.729Aでは、LPC情報を10bit、フレーム電力情報を5bitで量子化する。一方、AMRでは、LPC情報を29bit、フレーム電力情報を6bitで量子化する。ここで、LPC情報はLSPパラメータに変換して量子化される。このようにG.729AとAMRでは、量子化するためのビット割り当ても異なっている。図22(a), (b)はそれぞれG.729AとAMRにおける非音声符号(CN符号)構成図である。

【0021】G.729Aでは図22(a)に示すように非音声

符号のサイズは15bitであり、LSP符号I_LSPg(10bit)と電力符号I_POWg(5bit)で構成される。また、各符号はG.729Aの量子化器が持つ符号帳のインデックス(要素番号)で構成されており、詳細は以下の通りである。すなわち、(1)LSP符号I_LSPgは、符号L_{G1}(1bit)、L_{G2}(5bit)、L_{G3}(4bit)で構成され、L_{G1}は、LSP量子化器の予測係数の切り替え情報、L_{G2}、L_{G3}はLSP量子化器の符号帳CB_{G1}、CB_{G2}の各インデックス、(2)電力符号は、電力量子化器の符号帳CB_{G3}のインデックスである。AMRでは図22(b)に示すように非音声符号のサイズは35bitであり、LSP符号I_LSPa(29bit)と電力符号I_POWa(6bit)で構成される。また、各符号はAMRの量子化器が持つ符号帳のインデックスで構成されており、詳細は以下の通りである。すなわち、(1)LSP符号I_LSPaは、符号L_{A1}(3bit)、L_{A2}(8bit)、L_{A3}(9bit)、L_{A4}(9bit)で構成され、各符号は、LSP量子化器の符号帳GB_{A1}、GB_{A2}、GB_{A3}、GB_{A4}の各インデックス、(2)電力符号は、電力量子化器の符号帳GB_{A5}のインデックスである。

【0022】・DTX制御

次にDTXの制御方法について述べる。図23にG.729A、図24、図25にAMRのDTX制御の時間的流れを示す。まず、図23を参考にG.729AのDTX制御について説明する。G.729Aでは、VADが音声区間(VAD_flag=1)から非音声区間(VAD_flag=0)の変化を検出すると非音声区間の最初のフレームをSIDフレームとして設定する。SIDフレームは、上述した方法によるCN情報の生成、CN情報の量子化により作成され、受信側に伝送される。非音声区間では、フレーム毎に信号の変動を観測し、変動が検出されたフレームのみをSIDフレームとして設定し、再度CN情報の伝送を行う。変動なしと判定されたフレームは非伝送フレームとして設定し、情報の伝送は行わない。また、SIDフレーム間には最低非伝送フレームが2フレーム以上含まれるように制限している。変動の検出は、現フレームと最後に伝送したSIDフレームのCN情報の変化量を測定することにより行う。以上のように、G.729AではSIDフレームの設定が非音声信号の変動に対して適応的に行われる。

【0023】次に図24、図25を参考にAMRのDTX制御

について説明する。AMRでは、図24に示すようにSIDフレームの設定方法がG.729Aの適応制御と異なり基本的に8フレーム毎に定期的に設定される。ただし、長い音声区間後の非音声区間への変化点では、図25に示すようにハングオーバー制御を行う。具体的には、変化点以後7フレームが非音声区間(VAD_flag=0)にもかかわらず音声区間として設定され、通常の音声符号化処理が行われる。この区間をハングオーバーと呼ぶ。このハングオーバーは、最後にSIDフレームが設定されてからの経過フレーム数(P-FRM)が23フレーム以上の場合に設定される。これにより、変化点(非音声区間の始点)でのCN情報が音声区間(過去8フレーム)の特徴パラメータより求められるのを防止し、音声から非音声への変化点における音質を向上させることができるもの。

【0024】その後、8フレーム目が最初のSIDフレーム(SID_FIRSTフレーム)として設定され、SID_FIRSTフレームではCN情報の伝送は行わない。これはハングオーバー区間において受信側の復号器で復号信号からCN情報を生成できるためである。SID_FIRSTフレーム以後、3フレーム目がSID_UPDATEフレームとして設定され、ここで初めてCN情報の伝送が行われる。その後の非音声区間では、8フレーム毎にSID_UPDATEフレームが設定される。SID_UPDATEフレームは上述した方法により作成されて受信側へ伝送される。その他のフレームは非伝送フレームと設定されCN情報の伝送は行われない。

【0025】また、図24に示すように最後にSIDフレームが設定されてからの経過フレームが23フレーム以下の場合、ハングオーバー制御を行わない。この場合は、変化点のフレーム(非音声区間の最初のフレーム)がSID_UPDATEとして設定されるが、CN情報の算出を行わず最後に伝送したCN情報を再度伝送する。以上のようにAMRのDTX制御は、G.729Aのような適応制御を行わず固定制御でCN情報の伝送を行うため、音声から非音声への変化点を考慮して適宜ハングオーバー制御が行われる以上に示したようにG.729AとAMRの非音声圧縮機能は、基本原理は同じであるが、CN情報生成、量子化、DTX制御方法ともに異なっている。

【0026】

【発明が解決しようとする課題】従来技術1において、各通信システムが非音声圧縮機能を持つ場合の構成図を図26に示す。タンデム接続の場合、前述のように符号化方式1の音声符号を一旦再生信号に復号して符号化方式2により再度符号化を行う構成となる。各システムに非音声圧縮機能を持つ場合、図26のように符号変換部3のVAD部3cは符号化方式1によって符号／復号(情報圧縮)された再生信号を対象に音声／非音声区間の判定を行うことになる。このため、VAD部3cの音声／非音声区間の判定精度が低下し、誤判定による話頭切れ等の問題が生じ、音質が劣化する場合がある。このため、符号化方式2ではすべてを音声区間として処理するといった対策が考

えられるが、これでは最適な非音声圧縮が行えず本来の非音声圧縮による伝送効率向上の効果が損なわれる。更に、非音声区間では符号化方式1の復号器1aで生成されたコンフォートノイズから符号化方式2のCN情報を求める事になるため、入力信号に類似したノイズを発生させるためのCN情報としては必ずしも最適でない。

【0027】又、従来技術2は、従来技術1(タンデム接続)に比べ音質劣化と伝送遅延が少ない優れた音声符号変換方法であるが、非音声圧縮機能が考慮されていないという問題がある。つまり、従来技術2では入力される音声符号が常に音声区間として符号化された情報を想定しているため、非音声圧縮機能によりSIDフレーム又は非伝送フレームが生じた場合、正常な変換動作が行えない。

【0028】本発明の目的は、非音声符号化方法が異なる2つの音声通信システム間の通信において、送信側の非音声符号化方法で符号化したCN符号をCN信号に復号しなくても受信側の非音声符号化方法に応じたCN符号に変換することである。本発明の別の目的は、送信側と受信側のフレーム長の相違やDTX制御の相違を考慮して送信側のCN符号を受信側のCN符号に変換することである。本発明の別の目的は、非音声符号化方法や音声符号化方法が異なる2つの音声通信システム間の通信において、高品質な非音声符号変換及び音声符号変換を実現することである。

【0029】

【課題を解決するための手段】本発明の第1では、入力信号に含まれる非音声信号を第1の音声符号化式の非音声圧縮機能により符号化して得られた第1の非音声符号を一旦非音声信号に復号することなく第2の音声符号化方式の第2の非音声符号に変換する。例えば、第1の非音声符号を第1の複数の要素符号に分離し、第1の複数の要素符号を第2の非音声符号を構成する第2の複数の要素符号に変換し、この変換により得られた第2の複数の要素符号を多重化して第2の非音声符号を出力する。本発明によれば、非音声符号化方法が異なる2つの音声通信システム間の通信において、送信側の非音声符号化方法で符号化した非音声符号(CN符号)をCN信号に復号しなくても受信側の非音声符号化方法に応じた非音声符号(CN符号)に変換することができ、高品質な非音声符号変換を実現できる。

【0030】本発明の第2では、非音声区間の所定フレームにおいてのみ非音声符号を伝送し(非音声フレーム)、それ以外の非音声区間のフレーム(非伝送フレーム)では非音声符号を伝送せず、フレーム単位の符号情報に、音声フレーム、非音声フレーム、非伝送フレームの別を示すフレームタイプ情報を付加する。非音声符号の変換に際して、フレームタイプ情報に基いてどのフレームの符号であるか識別し、非音声フレーム、非伝送フレームの場合には、第1、第2の非音声符号化方式におけるフレーム長の

差、および非音声符号の伝送制御の相違を考慮して第1の非音声符号を第2の非音声符号に変換する。

【0031】例えば、(1)第1の非音声符号化方式が、非音声区間における所定フレーム数毎に平均した非音声符号を伝送すると共に、その他のフレームでは非音声符号を伝送しない方式であり、(2)第2の非音声符号化方式が、非音声区間における非音声信号の変化の度合が大きいフレームにおいてのみ非音声符号を伝送し、その他のフレームでは非音声符号を伝送せず、しかも、連続して非音声符号を伝送しない方式であり、更に、(3)第1の非音声符号化方式のフレーム長が、第2の非音声符号化方式のフレーム長の2倍であるとき、(a)第1の非音声符号化方式における非伝送フレームの符号情報を第2の非音声符号化方式における2つの非伝送フレームの符号情報に変換し、(b)第1の非音声符号化方式における非音声フレームの符号情報を、第2の非音声符号化方式における非音声フレームの符号情報と非伝送フレームの符号情報との2つのフレームに変換する。

【0032】又、音声区間から非音声区間に変化するとき、第1の非音声符号化方式が、変化点のフレームを含めて連続nフレームは音声フレームとみなして音声符号を伝送し、次のフレームは非音声符号を含まない最初の非音声フレームとしてフレームタイプ情報を伝送する場合、(a)第1の非音声符号化方式における前記最初の非音声フレームが検出されたとき、第1の音声符号化方式における直前n個の音声フレームの音声符号を逆量子化して得られる逆量子化値を平均化し、(b)平均値を量子化して前記第2の非音声符号化方式の非音声フレームにおける非音声符号を求める。

【0033】又、別の例として、(1)第1の非音声符号化方式が、非音声区間における非音声信号の変化の度合が大きいフレームにおいてのみ非音声符号を伝送し、その他のフレームでは非音声符号を伝送せず、また、連続して非音声符号を伝送しない方式であり、(2)第2の非音声符号化方式が、非音声区間における所定フレーム数N毎に平均した非音声符号を伝送すると共に、その他のフレームでは非音声符号を伝送しない方式であり、更に、(3)第1の非音声符号化方式のフレーム長が、第2の非音声符号化方式のフレーム長の半分であるとき、(a)第1の非音声符号化方式の連続する $2 \times N$ フレームにおける各非音声符号の逆量子化値を平均し、平均値を量子化して第2の非音声符号化方式におけるNフレーム毎の各フレームの非音声符号に変換し、(b) Nフレーム毎以外のフレームについては、第1の非音声符号化方式の連続する2つのフレームの符号情報をフレームタイプに関係なく第2の非音声符号化方式の1つの非伝送フレームの符号情報を変換する。

【0034】又、音声区間から非音声区間に変化するとき、前記第2の非音声符号化方式が、変化点のフレームを含めて連続nフレームは音声フレームとみなして音声

符号を伝送し、次のフレームは非音声符号を含まない最初の非音声フレームとしてフレームタイプ情報を伝送する場合、(a)第1の非音声フレームの非音声符号を逆量子化して複数の要素符号の逆量子化値を発生し、同時に、予め定めた、あるいはランダムな別の要素符号の逆量子化値を発生し、(b)連続する2フレームの各要素符号の逆量子化値を第2音声符号化方式の量子化テーブルを用いてそれぞれ量子化して第2音声符号化方式の1フレーム分の音声符号に変換し、(c) nフレーム分の第2音声符号化方式の音声符号を出力した後、非音声符号を含まない前記最初の非音声フレームのフレームタイプ情報を送出する。以上本発明の第2によれば、送信側と受信側のフレーム長の相違やDTX制御の相違を考慮して非音声信号に復号することなく送信側の非音声符号(CN符号)を受信側の非音声符号(CN符号)に変換することができ、高品質な非音声符号変換を実現できる。

【0035】

【発明の実施の形態】(A)本発明の原理

図1は本発明の原理説明図であり、符号化方式1と符号化方式2としてAMRやG.729AなどのCELP(Code ExcitedLinear Prediction)方式をベースとした符号化方式が用いられ、各符号化方式は前述した非音声圧縮機能を持つものとする。図1において、入力信号xinが符号化方式1の符号器51aへ入力すると、符号器51aは入力信号を符号化して符号データbst1を出力する。このとき、符号化方式1の符号器51aは非音声圧縮機能によりVAD部51bの判定結果(VAD_flag)に応じて音声・非音声区間の符号化処理を行う。従って、符号データbst1は音声符号か又は、CN符号で構成される。また、符号データbst1にはそのフレームが音声フレームであるかSIDフレームであるか(又は非伝送フレームであるか)を表すフレームタイプ情報Ftype1が含まれる。

【0036】フレームタイプ検出部52は、入力された符号データbst1からフレームタイプFtype1を検出し、変換制御部53へフレームタイプ情報Ftype1を出力する。変換制御部53は、フレームタイプ情報Ftype1に基いて音声区間、非音声区間を識別し、識別結果に応じて適切な変換処理を選択し、制御スイッチS1、S2の切り替えを行う。フレームタイプ情報Ftype1がSIDフレームであれば、非音声符号変換部60が選択される。非音声符号変換部60において、まず符号データbst1を符号分離部61に入力する。符号分離部61は符号データbst1を構成する符号化方式の1の要素CN符号に分離する。各要素CN符号はそれぞれCN符号変換部62₁～62nへ入力され、各CN符号変換部62₁～62nは各要素CN符号をCN情報に復号することなくそれぞれ符号化方式2の要素CN符号に直接変換する。符号多重部63は変換された各要素CN符号を多重化し、符号化方式2の非音声符号bst2として符号化方式2の復号器54へ入力する。

【0037】フレームタイプ情報Ftype1が非伝送フレー

ムの場合には変換処理を行わない。この場合、非音声符号bst2には非伝送フレームのフレームタイプ情報のみが含まれる。フレームタイプ情報Ftype1が音声フレームの場合には、従来技術1または従来技術2にしたがって構成した音声符号変換部70が選択される。音声符号変換部70は従来技術1または従来技術2にしたがって音声符号変換処理を行い、符号化方式2の音声符号で構成される符号データbst2が出力する。以上より、音声符号にフレームタイプ情報Ftype1を含ませたから、該情報を参照することによりフレームタイプを識別できる。このため、符号化方式変換部においてVAD部を不用にでき、しかも、音声区間と非音声区間の誤判定をなくすことができる。

【0038】又、符号化方式1のCN符号を一旦復号信号(CN信号)に戻さずに直接符号化方式2のCN符号に変換するため、受信側において入力信号に対して最適なCN情報を得ることができる。これにより、非音声圧縮機能による伝送効率の向上効果を損なうことなく、自然な背景雑音を再生することができる。また、音声フレームに加えSIDフレームおよび非伝送フレームに対しても正常な符号変換処理を行うことができる。これにより、非音声圧縮機能を持つ異なる音声符号化方式間での符号変換が可能となる。また、異なる非音声/音声圧縮機能を持つ2つの音声符号化方式間での符号変換が、非音声圧縮機能の伝送効率向上効果を維持しつつ、かつ、品質劣化と伝送遅延を抑えつつ、可能となるためその効果は大きい。

【0039】(B) 第1実施例

図2は本発明の非音声符号変換の第1実施例の構成図であり、符号化方式1としてAMRとしてG.729Aを用いた場合の例を示している。図2において、AMRの符号器(図示せず)より第nフレーム目の回線データすなわち音声符号bst1(n)が端子1に入力する。フレームタイプ検出部52は、回線データbst1(n)に含まれるフレームタイプ情報Ftype1(n)を抽出し変換制御部53に出力する。AMRのフレームタイプ情報Ftype1(n)は、音声フレーム(SPEECH)、SIDフレーム(SID_FIRST)、SIDフレーム(SID_UPDATE)、非伝送フレーム(NO_DATE)の4通りである(図24～図25参照)。非音声符号変換部60では、フレームタイプ情報Ftype1(n)に応じてCN符号変換制御を行う。

【0040】このCN符号変換制御では、AMRとG.729Aのフレーム長の違いを考慮する必要がある。図3に示すようにAMRのフレーム長は20msであり、これに対してG.729Aのフレーム長は10msである。したがって、変換処理はAMRの1フレーム(第nフレーム)をG.729Aの2フレーム(第m, m+1フレーム)として変換することになる。図4にAMRからG.729Aへのフレームタイプの変換制御手順を示す。以下に各場合について順に説明する。

【0041】(a) Ftype1(n)=SPEECHの場合

図4(a)に示すようにFtype1(n)=SPEECHの場合には、図2中の制御スイッチS1, S2が端子2に切り替えられ、音声

符号変換部70で符号変換処理が行われる。

(b) Ftype1(n)=SID_UPDATEの場合

次に、Ftype1(n)=SID_UPDATEの場合について説明する。図4(b-1)に示すようにAMRの1フレームがSID_UPDATEフレームである場合、G.729Aの第mフレームをSIDフレームと設定してCN符号変換処理を行う。すなわち、図2中のスイッチが端子3に切り替えられ、非音声符号変換部60は、AMRのCN符号bst1(n)をG.729Aの第mフレームのCN符号bst2(m)に変換する。また、図23で説明したようにG.729AではSIDフレームが続けて設定されることはないから、次フレームの第m+1フレームは非伝送フレームとして設定する。各CN要素符号変換部(LSP変換部62₁、フレーム電力変換部62₂)の動作について以下に説明する。

【0042】先ず、CN符号bst1(n)が符号分離部61に入力すれば、符号分離部61はCN符号bst1(n)をLSP符号I_LSP1(n)とフレーム電力符号I_POW1(n)に分離し、I_LSP1(n)をAMRと同じ量子化テーブルを持つLSP逆量子化器81に入力し、I_POW1(n)をAMRと同じ量子化テーブルを持つフレーム電力逆量子化器91に入力する。

【0043】LSP逆量子化器81は入力されたLSP符号I_LSP1(n)を逆量子化し、AMRのLSPパラメータLSP1(n)を出力する。すなわち、LSP逆量子化器81は逆量子化結果であるLSPパラメータLSP1(n)を、そのままG.729Aの第mフレームのLSPパラメータLSP2(m)としてLSP量子化器82へ入力する。LSP量子化器82はLSP2(m)を量子化し、G.729AのLSP符号I_LSP2(m)を出力する。ここでLSP量子化器82の量子化方法は任意であるが、使用する量子化テーブルはG.729Aで用いられているものと同じものである。

【0044】フレーム電力逆量子化器91は入力されたフレーム電力符号I_POW1(n)を逆量子化し、AMRのフレーム電力パラメータPOW1(n)を出力する。ここで、AMRとG.729Aのフレーム電力パラメータは、表1に示すようにAMRは入力信号領域、G.729AはLPC残差信号領域というようにフレーム電力を計算する際の信号領域が異なる。したがって、フレーム電力修正部92は、AMRのPOW1(n)をG.729Aで使用できるようにLSP残差信号領域に後述する手順に従って修正する。以上により、フレーム電力修正部92は、POW1(n)を入力としてG.729Aのフレーム電力パラメータPOW2(m)を出力する。フレーム電力量子化器93は、POW2(m)を量子化し、G.729Aのフレーム電力符号I_POW2(m)を出力する。ここでフレーム電力量子化器93の量子化方法は任意であるが、使用する量子化テーブルはG.729Aで用いられているものと同じものである。符号多重化部63はI_LSP2(m)とI_POW2(n)を多重化し、G.729AのCN符号bst2(m)として出力する。第m+1フレームは非伝送フレームとして設定されるため変換処理は行わない。したがって、bst2(m+1)には非伝送フレームを表すフレームタイプ情報のみが含まれる。

【0045】(c) Ftype1(n)=NO_DATAの場合

次にフレームタイプ情報Ftype1(n)=NO_DATAの場合は、

図4(c)のように第m、m+1フレームともに非伝送フレームとして設定される。この場合、変換処理は行わずbst2(m), bst2(m+1)には非伝送フレームを表すフレームタイプ情報のみが含まれる。

【0046】(d) フレーム電力修正法

$$E1 = \sqrt{\frac{1}{N_1} \sum_{n=0}^{N_1-1} err(n)^2} \quad (2)$$

である。err(n) (n=0, ..., N₁-1, N₁: G.729Aのフレーム長(80サンプル))はLPC残差信号であり、入力信号s(n) (n=0, ..., N₁-1)とs(n)から求めたLPC係数

$$err(n) = s(n) + \sum_{i=1}^{10} \alpha_i s(n-i) \quad (3)$$

により求められる。

【0048】一方、AMRの対数電力POW2は、次式を基に算出される。

$$E2 = \sqrt{\frac{1}{N_2} \sum_{n=0}^{N_2-1} s(n)^2} \quad (4)$$

である。また、N₂は、AMRのフレーム長(160サンプル)である。式(2)、式(5)から明らかのように、G.729AとAMRでは電力E1、E2を算出するのに各々残差err(n)、入力信号s(n)と異なる領域の信号を用いている。したがって、その間を変換する電力修正部が必要となる。修正方法は任意であるが、例えば以下の方法を考えられる。

【0049】・G.729AからAMRへの修正

図5(a)に処理フローを示す。まずG.729Aの対数電力POW1より電力E1を求める。

$$E1 = 10^{(POW1/20)} \quad (6)$$

$$d_s(n) = d_err(n) - \sum_{i=1}^{10} \alpha_i d_s(n-i) \quad (7)$$

ここで、 α_i (i=1, ..., 10)はLSP逆量化値から求められたG.729AのLPC係数である。またd_s(-i) (i=1, ..., 10)の初期値は0とする。d_s(n)の電力を

$$POW2 = \log_2 \sqrt{\frac{1}{N_1} \sum_{n=0}^{N_1-1} d_s(n)^2} \quad (8)$$

【0051】・AMRからG.729Aへの修正

図5(b)に処理フローを示す。まず、AMRの対数電力POW2より電力E2を求める。

$$E2 = 2^{POW2} \quad (10)$$

電力がE2となる擬似入力信号d_s(n) (n=0, ..., N₂-1)を次式より生成する。

$$d_s(n) = E2 \cdot q(n) \quad (11)$$

$$d_err(n) = d_s(n) + \sum_{i=1}^{10} \alpha_i d_s(n-i) \quad (12)$$

G.729Aの対数電力POW1は、次式を基に算出される。

$$POW1 = 20 \log_{10} E1 \quad (1)$$

ここで、

【数1】

α_i (i=1, ..., 10)を用いて次式

【0047】

【数2】

$$POW2 = \log_2 E2 \quad (4)$$

ここで、

【数3】

$$d_err(n) = E1 \cdot q(n) \quad (7)$$

ここで、q(n) (n=0, ..., N₁-1)は、電力が1に正規化されたランダムノイズ信号である。d_err(n)をLPC合成フィルタに通して、擬似信号(入力信号領域)d_s(n) (n=0, ..., N₁-1)を生成する。

【0050】

【数4】

算出し、AMRの電力E2として用いる。したがって、AMRの対数電力POW2は、次式で求められる。

【数5】

$$POW2 = \log_2 \sqrt{\frac{1}{N_2} \sum_{n=0}^{N_2-1} d_s(n)^2} \quad (9)$$

ここで、q(n)は、電力が1に正規化されたランダムノイズ信号である。d_s(n)をLPC逆合成フィルタに通して、擬似信号(LPC残差信号領域)d_err(n) (n=0, ..., N₂-1)を生成する。

【0052】

【数6】

ここで、 α_i ($i=1, \dots, 10$) は LSP 逆量子化値から求められた AMR の LPC 係数である。また、 $d_s(-i)$ ($i=1, \dots, 10$) の初期値は 0 とする。 $d_{err}(n)$ の電力

$$POW1 = 20 \log_{10} \sqrt{\frac{1}{N_2} \sum_{n=0}^{N_2-1} d_{err}(n)^2} \quad (13)$$

で求められる。

【0053】(e) 第1実施例の効果

以上説明した通り、第1実施例によれば AMR の CN 符号である LSP 符号とフレーム電力符号を G.729A の CN 符号に直接変換できる。また、音声符号変換部 70 と非音声符号変換部 60 を切り替えることにより非音声圧縮機能を備えた AMR から符号データ（音声符号、非音声符号）を一旦再生音声に復号することなしに非音声圧縮機能を備えた G.729A の符号データに正常に変換することができる。

【0054】(C) 第2実施例

図 6 は本発明の第2実施例の構成図であり、図 2 の第1実施例と同一部分には同一符号を付している。第2実施例は、第1実施例と同様に符号化方式 1 として AMR として G.729A を用いた場合において、フレームタイプ検出部 52 で検出した AMR のフレームタイプが $Ftype1(n)=SID_FIRST$ の場合の変換処理を実現するものである。図 4 の (b-2) で示すように AMR の 1 フレームが SID_FIRST フレームの場合も、第1実施例の SID_UPDATE フレームの場合（図 4 の (b-1)）と同様に G.729A の第 m フレームを SID フレーム、第 $m+1$ フレームを非伝送フレームと設定して変換処理を行える。しかし、図 2 で説明したように AMR の SID_FIRST フレームでは、ハングオーバー制御により CN 符号が伝送されてきていないことを考慮する必要がある。すなわち、図 2 の第1実施例の構成では、bst1(n) が送られてこないためこのままでは G.729A の CN パラメータである LSP2(m) と POW2(m) を求めることができない。

【0055】そこで、第2実施例では、SID_FIRST フレーム直前に伝送された過去 7 フレームの音声フレームの情報を用いてこれらを算出する。以下に変換処理について説明する。上述の通り、SID_FIRST フレームにおける LSP2(m) は、音声符号変換部 70 における LSP 符号変換部 4b の LSP 逆量子化部 4b₁（図 17 参照）から出力する過去 7 フレーム分の LSP パラメータ OLD_LSP(1), (1=n-1, n-7) の平均値として算出する。したがって LSP バッファ部 83 は現フレームに対して常に過去 7 フレームの LSP パラメータを保持し、LSP 平均値算出部 84 は過去 7 フレーム分の LSP パラメータ OLD_LSP(1), (1=n-1, n-7) の平均値を算出して保持する。

【0056】POW2(m) も同様に過去 7 フレームのフレーム電力 OLD_POW(1), (1=n-1, n-7) の平均値として算出する。OLD_POW(1) は、音声符号変換部 70 におけるゲイン符号変換部 4e（図 17 参照）で生成される音源信号 EX(1) のフレーム電力として求められる。したがって、電力計算部 94 は音源信号 EX(1) のフレーム電力を計算し、フレーム電力

を算出し、G.729A の電力 E1 として用いる。したがって、G.729A の対数電力 POW1 は、次式

【数 7】

バッファ部 95 は、現フレームに対して常に過去 7 フレームのフレーム電力 OLD_POW(1) を保持し、電力平均値算出部 96 は過去 7 フレーム分のフレーム電力 OLD_POW(1) の平均値を算出して保持する。LSP 量子化器 82 及びフレーム電力量子化器 93 は、非音声区間においてフレームタイプが SID_FIRST でなければ、変換制御部 53 よりその旨が通知されるから、LSP 逆量子化器 81 及びフレーム電力逆量子化器 91 から出力する LSP パラメータ、フレーム電力パラメータを用いて G.729A の LSP 符号 I_LSP2(m) 及びフレーム電力符号 I_POW2(m) を求めて出力する。

【0057】しかし、非音声区間においてフレームタイプが SID_FIRST であれば、すなわち、 $Ftype1(n)=SID_FIRST$ であれば、変換制御部 53 よりその旨が通知される。これにより、LSP 量子化器 82 及びフレーム電力量子化器 93 は、LSP 平均値算出部 84 及び電力平均値算出部 96 で保持されている過去 7 フレーム分の平均 LSP パラメータ、平均フレーム電力パラメータを用いて G.729A の LSP 符号 I_LSP2(m) 及びフレーム電力符号 I_POW2(m) を求めて出力する。符号多重部 63 は、LSP 符号 I_LSP2(m) 及びフレーム電力符号 I_POW2(m) を多重化し、bst2(m) として出力する。また、第 $m+1$ フレームでは変換処理は行わず、bst2(m+1) には非伝送フレームを表すフレームタイプ情報のみを含めて送出する。

【0058】以上説明した通り、第2実施例によれば AMR のハングオーバー制御により変換するべき CN 符号が得られない場合でも、過去の音声フレームの音声パラメータを利用して CN パラメータを求め、G.729A の CN 符号を生成することができる。

【0059】(D) 第3実施例

図 7 に本発明の第3実施例の構成図を示し、第1実施例と同一部分には同一符号を付している。第3実施例は、符号化方式 1 として G.729A として AMR を用いた場合の例を示している。図 7 において、G.729A の符号器（図示せず）より第 m フレーム目の回線データ すなわち 音声符号 bst1(m) が 端子 1 に入力する。フレームタイプ検出部 52 は、bst1(m) に含まれるフレームタイプ $Ftype(m)$ を抽出し 変換制御部 53 に出力する。G.729A の $Ftype(m)$ は 音声フレーム (SPEECH)、SID フレーム (SID)、非伝送フレーム (NO_DA TA) の 3 通りである（図 23 参照）。変換制御部 53 は フレームタイプに基いて音声区間、非音声区間を識別して制御スイッチ S1, S2 を切り替える。

【0060】非音声符号変換部 60 は、非音声区間においてフレームタイプ情報 $Ftype(m)$ に応じて CN 符号変換処理の制御を行う。ここで、第1実施例と同様に AMR と G.72

9Aのフレーム長の違いを考慮する必要がある。すなわち、G.729Aの2フレーム分(第m, 第m+1フレーム)をAMRの1フレーム分(第nフレーム)として変換することになる。また、G.729AからAMRへの変換では、DTX制御の相違点を考慮して変換処理を制御する必要がある。

【0061】図8に示すように、Ftype1(m), Ftype1(m+1)がともに音声フレーム(SPEECH)の場合には、AMRの第nフレームも音声フレームとして設定する。すなわち、図7の制御スイッチS1, S2が端子2, 4に切り替えられ、音声符号変換部70が従来技術2にしたがって音声符号の符号変換処理を行う。また、図9に示すようにFtype1(m), Ftype1(m+1)が共に非伝送フレーム(NO_DATA)の場合には、AMRの第nフレームも非伝送フレームに設定し、変換処理は行わない。すなわち、図7の制御スイッチS1, S2が端子3, 5に切り替えられ、符号多重部63は非伝送フレームのフレームタイプ情報のみを送出する。従って、bst2(n)には非伝送フレームを表すフレームタイプ情報のみが含まれる。

【0062】次に、図10に示すような非音声区間でのCN符号の変換方法について説明する。図10は非音声区間でのCN符号変換方法の時間的な流れを示す。非音声区間において、図7のスイッチS1, S2は端子3, 5に切り替えられ、非音声符号変換部60がCN符号の変換処理を行う。この変換処理において、G.729AとAMRのDTX制御の相違点を考慮する必要がある。G.729AにおけるSIDフレームの伝送制御は適応的であり、CN情報(非音声信号)の変動に応じてSIDフレームが不定期に設定される。一方、AMRではSIDフレーム(SID_UPDATA)は8フレーム毎に定期的に設定されるようになっている。したがって、非音声区間では図10に示すように変換元のG.729Aのフレームタイプ(SID or NO_DATA)に関係なく、変換先のAMRに合わせて8フレーム毎(G.729Aで16フレームに相当)にSIDフレーム(SID_UPDATA)へ変換する。また、その他の7フレームは非伝送区間(NO_DATA)となるように変換を行う。

【0063】具体的には、図10中のAMRの第nフレームにおけるSID_UPDATAフレームへの変換では、現フレーム(第m, 第m+1フレーム)を含む過去16フレーム(第m-14, …, 第m+1)(AMRでは8フレームに相当)の間に受信したSIDフレームのCNパラメータから平均値を求め、AMRのSID_UPDATAフレームのCNパラメータへ変換する。図7を参考に変換処理について説明する。

【0064】第kフレームでG.729AのSIDフレームが受信されると、符号分離部61はCN符号bst1(k)をLSP符号I_LSP1(k)とフレーム電力符号I_POW1(k)に分離し、I_LSP1(k)をG.729Aと同じ量子化テーブルを持つLSP逆量子化器81に入力し、I_POW1(k)をG.729Aと同じ量子化テーブルを持つフレーム電力逆量子化器91に入力する。LSP逆量子化器81はLSP符号I_LSP1(k)を逆量子化してG.729AのLSPパラメータLSP1(k)を出力する。フレーム電力逆量子化器91はフレーム電力符号I_POW1(k)を逆量子化してG.729Aのフレーム電力パラメータPOW1(k)を出力する。

【0065】G.729AとAMRのフレーム電力パラメータは、表1に示したようにG.729AはLPC残差信号領域、AMRは入力信号領域というようにフレーム電力を計算する際の信号領域が異なる。したがって、フレーム電力修正部92はG.729AのLSP残差信号領域のパラメータPOW1(k)をAMRで使用できるように入力信号領域に修正する。この結果、フレーム電力修正部92はPOW2(k)を入力されてAMRのフレーム電力パラメータPOW2(k)を出力する。求められたLSP(k), POW2(k)は、それぞれバッファ部85, 97に入力される。ここでk=m-14, …, m+1であり、過去16フレームで受信したSIDフレームの各CNパラメータがバッファ部85, 97で保持される。ここで、もし過去16フレームにおいて受信したSIDフレームが無い場合には、最後に受信したSIDフレームのCNパラメータを用いる。

【0066】平均値算出部86, 98はバッファ保持データの平均値を算出し、AMRのCNパラメータLSP2(n), POW2(n)として出力する。LSP量子化器82はLSP2(n)を量子化し、AMRのLSP符号I_LSP2(n)を出力する。ここでLSP量子化器82の量子化方法は任意であるが、使用する量子化テーブルはAMRで用いられているものと同じものである。フレーム電力量化器93はPOW2(n)を量子化し、AMRのフレーム電力符号I_POW2(n)を出力する。ここでフレーム電力量化器93の量子化方法は任意であるが、使用する量子化テーブルはAMRで用いられているものと同じものである。符号多重化部63はI_LSP2(n)とI_POW2(n)を多重化すると共にフレームタイプ情報(=U)を付加してbst2(n)として出力する。

【0067】以上説明した通り、第3実施例によれば非音声区間において変換元のG.729Aのフレームタイプに関わらず、CN符号の変換処理を変換先のAMRのDTX制御に合わせて定期的に行う場合、変換処理が行われるまでに受信したG.729AのCNパラメータの平均値をAMRのCNパラメータとして用いることでAMRのCN符号を生成することができる。また、音声符号変換部とCN符号変換部を切り替えることにより非音声圧縮機能を備えたG.729Aの符号データ(音声符号、非音声符号)を一旦再生音声に復号することなしに非音声圧縮機能を備えたAMRの符号データに正常に変換することができる。

【0068】(E) 第4実施例

図11は本発明の第4実施例の構成図であり、図7の第3実施例と同一部分には同一符号を付している。図12は第4実施例における音声符号変換部70の構成図である。第4実施例は、第3実施例と同様に符号化方式1としてG.729A 2としてAMRを用いた場合において、音声区間から非音声区間への変化点でのCN符号変換処理を実現するものである。図13に変換制御方法の時間的な流れを示す。G.729Aの第mフレームが音声フレーム、第m+1フレームがSIDフレームである場合、そこは音声区間から非音声区間への変化点である。AMRではこのような変化点でハングオーバー

バー制御を行う。なお、最後にSID_UPDATAフレームへ変換処理が行われてから区間変更フレームまでのAMRにおける経過フレーム数が23フレーム以下の場合には、ハングオーバー制御は行われない。以下では、経過フレームが23フレームより大きく、ハングオーバー制御を行う場合について説明する。

【0069】ハングオーバー制御を行う場合、変換点フレームから7フレーム(第n, …, 第n+7フレーム)は非音声フレームにもかかわらず、音声フレームとして設定する必要がある。従って、図13(a)に示すようにG.729Aの第m+1フレーム～第m+13フレームは、非音声フレーム(SIDフレーム or 非伝送フレーム)にもかかわらず、変換先のAMRのDTX制御に合わせて音声フレームとみなして変換処理を行う。以下、図11、図12を参考に変換処理について説明する。

【0070】音声区間から非音声区間への変換点において、G.729AからAMRの音声フレームに変換するためには、音声符号変換部70を用いて変換処理するしかない。しかし、変換点以降ではG.729A側が非音声フレームであるため、このままでは音声符号変換部70の入力となるG.729Aの音声パラメータ(LSP、ピッチラグ、代数符号、ピッチゲイン、代数符号ゲイン)を得ることができない。そこで、図12に示すようにLSPと代数符号ゲインは、非音声符号変換部60で最後に受信したCNパラメータLSP1(k), P0_W1(k) ($k < n$)で代用し、その他のパラメータ(ピッチラグlag(m), ピッチゲインGa(m), 代数符号code(m))については、ピッチラグ生成部101、代数符号生成部102、ピッチゲイン生成部103で聴覚的に悪影響の無い程度で任意に生成する。生成方法はランダムに生成しても、固定値により生成してもよい。ただし、ピッチゲインについては最小値(0.2)を設定することが望ましい。

【0071】音声区間及び音声→非音声区間への切り替わり時、音声符号変換部70は以下のように動作する。音声区間において、符号分離部71は入力するG.729Aの音声符号より、LSP符号I_LSP1(m)、ピッチラグ符号I_LAG1(m)、代数符号I_CODE1(m)、ゲイン符号I_GAIN1(m)を分離し、それぞれLSP逆量子化器72a、ピッチラグ逆量子化器73a、代数符号逆量子化器74a、ゲイン逆量子化器75aに入力する。又、音声区間ににおいて、切換部77a～77eは変換制御部53からの指示により、LSP逆量子化器72a、ピッチラグ逆量子化器73a、代数符号逆量子化器74a、ゲイン逆量子化器75aの出力を選択する。

【0072】LSP逆量子化器72aは、G.729AのLSP符号を逆量子化してLSP逆量子化値を出力し、LSP量子化器72bは該LSP逆量子化値をAMRのLSP量子化テーブルを用いて量子化してLSP符号I_LSP2(n)を出力する。ピッチラグ逆量子化器73aは、G.729Aのピッチラグ符号を逆量子化してピッチラグ逆量子化値を出力し、ピッチラグ量子化器73bは該ピッチラグ逆量子化値をAMRのピッチラグ量子化テーブルを用いて量子化してピッチラグ符号I_LAG2(n)

を出力する。代数符号逆量子化器74aは、G.729Aの代数符号を逆量子化して代数符号逆量子化値を出力し、代数符号量子化器74bは該代数符号逆量子化値をAMRの代数符号量子化テーブルを用いて量子化して代数符号I_CODE2(n)を出力する。ゲイン逆量子化器75aは、G.729Aのゲイン符号を逆量子化してピッチゲイン逆量子化値Gaと代数ゲイン逆量子化値Gcを出力し、ピッチゲイン量子化器75bは該ピッチゲイン逆量子化値GaをAMRのピッチゲイン量子化テーブルを用いて量子化してピッチゲイン符号I_GAIN2a(n)を出力する。また、代数ゲイン量子化器75cは代数ゲイン逆量子化値GcをAMRのゲイン量子化テーブルを用いて量子化して代数ゲイン符号I_GAIN2c(n)を出力する。

【0073】符号多重化部76は、各量子化器72b～75b, 75cから出力するLSP符号、ピッチラグ符号、代数符号、ピッチゲイン符号、代数ゲイン符号を多重し、フレームタイプ情報(=S)を付加してAMRによる音声符号を作成して送出する。音声区間においては、以上の動作が繰り返され、G.729Aの音声符号をAMRの音声符号に変換して出力する。一方、音声→非音声区間への切り替わり時においてハングオーバー制御を行うものとすれば、切換部77aは変換制御部53からの指示に従って、非音声符号変換部60で最後に受信したLSP符号より得られたLSPパラメータLSP1(k)を選択してLSP量子化器72bに入力する。また、切換部77bはピッチラグ生成部101から発生するピッチラグパラメータlag(m)を選択してピッチラグ量子化器73bに入力する。また、切換部77cは代数符号生成部102から発生する代数符号パラメータcode(m)を選択して代数符号量子化器74bに入力する。また、切換部77dはピッチゲイン生成部103から発生するピッチゲインパラメータGa(m)を選択してピッチゲイン量子化器75bに入力する。また、切換部77eは非音声符号変換部60で最後に受信したフレーム電力符号I_POW1(k)より得られたフレーム電力パラメータPOW1(k)を選択して代数ゲイン量子化器75cに入力する。

【0074】LSP量子化器72bは切換部77aを介して非音声符号変換部60より入力したLSPパラメータLSP1(k)をAMRのLSP量子化テーブルを用いて量子化してLSP符号I_LSP2(n)を出力する。ピッチラグ量子化器73bは切換部77bを介してピッチラグ生成部101より入力したピッチラグパラメータをAMRのピッチラグ量子化テーブルを用いて量子化してピッチラグ符号I_LAG2(n)を出力する。代数符号量子化器74bは切換部77cを介して代数符号生成部102より入力した代数符号パラメータをAMRの代数符号量子化テーブルを用いて量子化して代数符号I_CODE2(n)を出力する。ピッチゲイン量子化器75bは切換部77dを介してピッチゲイン生成部103より入力したピッチゲインパラメータをAMRのピッチゲイン量子化テーブルを用いて量子化してピッチゲイン符号I_GAIN2a(n)を出力する。また、代数ゲイン量子化器75cは切換部77eを介し

て非音声符号変換部60より入力したフレーム電力パラメータPOW1(k)をAMRの代数ゲイン量子化テーブルを用いて量子化して代数ゲイン符号I_GAIN2c(n)を出力する。

【0075】符号多重化部76は、各量子化器72b～75b, 75cから出力するLSP符号、ピッチラグ符号、代数符号、ピッチゲイン符号、代数ゲイン符号を多重し、フレームタイプ情報(=S)を付加してAMRによる音声符号を作成して送出する。音声区間～非音成区間への変化点において、音声符号変換部70はAMRの7フレーム分の音声符号を送出するまで以上の動作を繰り返し、7フレーム分の音声符号の送出が完了すれば次の音声区間が検出されるまで音声符号の出力を停止する。

【0076】7フレーム分の音声符号の送出が完了すれば、交換制御部53の制御で図11のスイッチS1, S2が端子3, 5側に切り替わり、以後、非音声符号変換部60によるCN符号変換処理が行われる。図13(a)に示すようにハングオーバー後の第m+14, 第m+15フレーム(AMR側の第n+7フレーム)は、AMRのDTX制御に合わせてSID_FIRSTフレームとして設定する必要がある。ただし、CNパラメータの伝送は必要なく、したがって、符号多重部63はSID_FIRSTのフレームタイプを表す情報のみをbst2(m+7)に含めて出力する。以後、図7の第3実施例と同様にCN符号変換を行う。

【0077】以上は、ハングオーバー制御を行う場合におけるCN符号変換であるが、最後にSID_UPDATAフレームへ変換処理が行われてから変化点フレームまでのAMRにおける経過フレーム数が23フレーム以下の場合には、ハングオーバー制御は行われない。かかるハングオーバー制御を行わない場合の制御方法を図13(b)に示す。音声区間と非音声区間の境界フレームである第m, 第m+1フレームは、ハングオーバー時と同じように音声符号変換部70でAMRの音声フレームに変換して出力する。

【0078】次の第m+2, 第m+3フレームは、SID_UPDATAフレームに変換する。また、第m+4フレーム以後のフレームは第3実施例で述べた非音声区間における変換方法と同じ方法を用いる。次に非音声区間から音声区間への変化点でのCN符号変換方法について説明する。図14に変換制御方法の時間的な流れを示す。G.729Aの第mフレームが非音声フレーム(SIDフレーム or 非伝送フレーム)、第m+1フレームが音声フレームである場合、そこは非音声区間から音声区間への変化点である。この場合、音声の話頭切れ(音声の立ち上がりが消えてしまう)を防ぐため、AMRの第nフレームは音声フレームとして変換する。したがって、G.729Aの第mフレームは非音声フレームを音声フレームとして変換する。変換方法は、ハングオーバー時と同じように音声符号変換部70でAMRの音声フレームに変換して出力する。

【0079】以上説明した通り、本実施例によれば音声区間から非音声区間への変化点においてG.729Aの非音声フレームをAMRの音声フレームに変換する必要がある場合、G.729AのCNパラメータをAMRの音声パラメータとし

て代用してAMRの音声符号を生成することができる。

【0080】・付記

(付記1) 入力信号を第1の音声符号化方式で符号化して得られる第1の音声符号を、第2の音声符号化方式の第2の音声符号に変換する音声符号変換方法において、入力信号に含まれる非音声信号を第1の音声符号化方式の非音声圧縮機能により符号化して得られた第1の非音声符号を一旦非音声信号に復号することなく第2の音声符号化方式の第2の非音声符号に変換する、することを特徴とする音声符号変換方法。

【0081】(付記2) 入力信号を第1の音声符号化方式で符号化して得られる第1の音声符号を、第2の音声符号化方式の第2の音声符号に変換する音声符号変換方法において、入力信号に含まれる非音声信号を第1の音声符号化方式の非音声圧縮機能により符号化して得られた第1の非音声符号を第1の複数の要素符号に分離し、第1の複数の要素符号を前記第2の非音声符号を構成する第2の複数の要素符号に変換し、前記変換により得られた第2の複数の要素符号を多重化して第2の非音声符号を出力する、ことを特徴とする音声符号変換方法。

【0082】(付記3) 前記第1の要素符号は、非音声信号を一定サンプル数からなるフレームに分割し、フレーム毎に分析して得られる非音声信号の特徴を表す特徴パラメータを第1の音声符号化方式独自の量子化テーブルを用いて量子化して得られる符号であり、前記第2の要素符号は、前記特徴パラメータを第2の音声符号化方式独自の量子化テーブルを用いて量子化して得られる符号である、ことを特徴とする付記2記載の音声符号変換方法。

(付記4) 前記特徴パラメータは、非音声信号の周波数特性の概形を表わすLPC係数(線形予測係数)と非音声信号の振幅特性を表わすフレーム信号電力である、ことを特徴とする付記3記載の音声符号変換方法。

(付記5) 前記変換ステップにおいて、前記第1の複数の要素符号を第1の音声符号化方式と同じ量子化テーブルを持つ逆量子化器で逆量子化し、逆量子化により得られた複数の要素符号の逆量子化値を第2の音声符号化方式と同じ量子化テーブルを持つ量子化器で量子化して第2の複数の要素符号に変換する、ことを特徴とする付記2または付記3または4記載の音声符号変換方法。

【0083】(付記6) 入力信号の一定サンプル数をフレームとし、フレーム単位で音声区間における音声信号を第1の音声符号化方式で符号化して得られる第1の音声符号と、非音声区間における非音声信号を第1の音声符号化方式で符号化して得られる第1の非音声符号を混在して送信側より伝送し、これら第1の音声符号と第1の非音声符号をそれぞれ、第2の音声符号化方式による第2の音声符号と第2の非音声符号化方式による第2の非音声符号とにそれぞれ変換し、変換により得られた第2の音声符号と第2の非音声符号を混在して受信側に伝送

する音声通信システムにおける音声符号変換方法において、非音声区間では所定のフレームにおいてのみ非音声符号を伝送し、それ以外のフレームでは非音声符号を伝送せず、前記フレーム単位の符号情報に、音声フレーム、非音声フレーム、符号を伝送しない非伝送フレームの別を示すフレームタイプ情報を附加し、フレームタイプ情報に基いてどのフレームの符号であるか識別し、非音声フレーム、非伝送フレームの場合には、第1、第2の非音声符号化方式におけるフレーム長の差、および非音声符号の伝送制御の相違を考慮して第1の非音声符号を第2の非音声符号に変換することを特徴とする音声符号変換方法。

【0084】(付記7) (1) 第1の非音声符号化方式が、非音声区間ににおける所定フレーム数毎に平均した非音声符号を伝送すると共に、その他のフレームでは非音声符号を伝送しない方式であり、(2) 第2の非音声符号化方式が、非音声区間ににおける非音声信号の変化の度合が大きいフレームにおいてのみ非音声符号を伝送し、その他のフレームでは非音声符号を伝送せず、しかも、連続して非音声符号を伝送しない方式であり、更に、(3) 第1の非音声符号化方式のフレーム長が、第2の非音声符号化方式のフレーム長の2倍であるとき、第1の非音声符号化方式における非伝送フレームの符号情報を第2の非音声符号化方式における2つの非伝送フレームの符号情報に変換し、第1の非音声符号化方式における非音声フレームの符号情報を、第2の非音声符号化方式における非音声フレームの符号情報と非伝送フレームの符号情報との2つに変換することを特徴とする付記6記載の音声符号変換方法。

【0085】(付記8) 音声区間から非音声区間に変化するとき、前記第1の非音声符号化方式が、変化点のフレームを含めて連続nフレームは音声フレームとみなして音声符号を伝送し、次のフレームは非音声符号を含まない最初の非音声フレームとしてフレームタイプ情報を伝送する場合、第1の非音声符号化方式における前記最初の非音声フレームが検出された時、第1の音声符号化方式における直前n個の音声フレームの音声符号を逆量子化して得られる逆量子化値を平均化し、平均値を量子化して前記第2の非音声符号化方式の非音声フレームにおける非音声符号を求める、ことを特徴とする付記7記載の音声符号変換方法。

【0086】(付記9) (1) 第1の非音声符号化方式が、非音声区間ににおける非音声信号の変化の度合が大きいフレームにおいてのみ非音声符号を伝送し、その他のフレームでは非音声符号を伝送せず、また、連続して非音声符号を伝送しない方式であり、(2) 第2の非音声符号化方式が、非音声区間ににおける所定フレーム数N毎に平均した非音声符号を伝送すると共に、その他のフレームでは非音声符号を伝送しない方式であり、更に、(3) 第1の非音声符号化方式のフレーム長が、第2の非音声符

号化方式のフレーム長の半分であるとき、第1の非音声符号化方式の連続する $2 \times N$ フレームにおける各非音声符号の逆量子化値を平均し、平均値を逆量子化して第2の非音声符号化方式におけるNフレーム毎のフレームの非音声符号とし、Nフレーム毎以外のフレームについては、第1の非音声符号化方式の連続する2つのフレームの符号情報をフレームタイプに関係なく第2の非音声符号化方式の1つの非伝送フレームの符号情報を変換することを特徴とする付記6記載の音声符号変換方法。

【0087】(付記10) 音声区間から非音声区間に変化するとき、前記第2の非音声符号化方式が、変化点のフレームを含めて連続nフレームは音声フレームとみなして音声符号を伝送し、次のフレームは非音声符号を含まない最初の非音声フレームとしてフレームタイプ情報を伝送する場合、第1の非音声フレームの非音声符号を逆量子化して複数の要素符号の逆量子化値を発生し、同時に、予め定めた、あるいはランダムな別の要素符号の逆量子化値を発生し、連続する2フレームの各要素符号の逆量子化値を第2音声符号化方式の量子化テーブルを用いてそれぞれ量子化して第2音声符号化方式の1フレーム分の音声符号に変換し、nフレーム分の第2音声符号化方式の音声符号を出力した後、非音声符号を含まない前記最初の非音声フレームのフレームタイプ情報を送出する、ことを特徴とする付記9記載の音声符号変換方法。

【0088】(付記11) 入力信号を第1の音声符号化方式で符号化して得られる第1の音声符号を、第2の音声符号化方式の第2の音声符号に変換する音声符号変換装置において、入力信号に含まれる非音声信号を第1の音声符号化方式の非音声圧縮機能により符号化して得られた第1の非音声符号を第1の複数の要素符号に分離する符号分離部、第1の複数の要素符号を、前記第2の非音声符号を構成する第2の複数の要素符号に変換する要素符号変換部、前記変換により得られた第2の各要素符号を多重化して第2の非音声符号を出力する符号多重部、を備えたことを特徴とする音声符号変換装置。

【0089】(付記12) 前記第1の要素符号は、非音声信号を一定サンプル数からなるフレームに分割し、フレーム毎に分析して得られる非音声信号の特徴を表す特徴パラメータを第1の音声符号化方式独自の量子化テーブルを用いて量子化して得られる符号であり、前記第2の要素符号は、前記特徴パラメータを第2の音声符号化方式独自の量子化テーブルを用いて量子化して得られる符号である、ことを特徴とする付記11記載の音声符号変換装置。

(付記13) 前記要素符号変換部は、前記第1の各要素符号を第1の音声符号化方式と同じ量子化テーブルに基いて逆量子化する逆量子化器、前記逆量子化により得られた各要素符号の逆量子化値を第2の音声符号化方式と同じ量子化テーブルに基いて量子化して第2の各要素符号に変換する量子化器、を備えたことを特徴とする付

記11または12記載の音声符号変換装置。

【0090】(付記14) 入力信号の一定サンプル数をフレームとし、フレーム単位で音声区間における音声信号を第1の音声符号化方式で符号化して得られる第1の音声符号と、非音声区間における非音声信号を第1の非音声符号化方式で符号化して得られる第1の非音声符号を混在して送信側より伝送し、これら第1の音声符号と第1の非音声符号をそれぞれ、第2の音声符号化方式による第2の音声符号と第2の非音声符号化方式による第2の非音声符号とにそれぞれ変換し、変換により得られた第2の音声符号と第2の非音声符号を受信側に伝送する音声通信システムにおける音声符号変換装置において、符号情報に付加されているフレームタイプ情報に基いて、音声フレーム、非音声フレーム、非音声区間において非音声符号を伝送しない非伝送フレームの別を識別するフレームタイプ識別部、非音声フレームにおける第1の非音声符号を、第1の非音声符号化方式と同じ量子化テーブルに基いて逆量子化し、得られた逆量子化値を第2の非音声符号化方式と同じ量子化テーブルに基いて量子化して第2の非音声符号に変換する非音声符号変換部、第1、第2の非音声符号化方式におけるフレーム長の差、および非音声符号の伝送制御の相違を考慮して前記非音声符号変換部を制御する変換制御部、を有することを特徴とする音声符号変換装置。

【0091】(付記15) (1)第1の非音声符号化方式が、非音声区間における所定フレーム数毎に平均した非音声符号を伝送すると共に、その他のフレームでは非音声符号を伝送しない方式であり、(2)第2の非音声符号化方式が、非音声区間における非音声信号の変化の度合が大きいフレームにおいてのみ非音声符号を伝送し、その他のフレームでは非音声符号を伝送せず、しかも、連続して非音声符号を伝送しない方式であり、更に、(3)第1の非音声符号化方式のフレーム長が、第2の非音声符号化方式のフレーム長の2倍であるとき、前記非音声符号変換部は、第1の非音声符号化方式における非伝送フレームの符号情報を第2の非音声符号化方式における2つの非伝送フレームの符号情報に変換し、第1の非音声符号化方式における非音声フレームの符号情報を、第2の非音声符号化方式における非音声フレームの符号情報と非伝送フレームの符号情報の2つに変換することを特徴とする付記14記載の音声符号変換装置。

【0092】(付記16) 音声区間から非音声区間に変化するとき、前記第1の非音声符号化方式が、変化点のフレームを含めて連続nフレームは音声フレームとみなして音声符号を伝送し、次のフレームは非音声符号を含まない最初の非音声フレームとしてフレームタイプ情報を伝送する場合、前記非音声符号変換部は、第1の音声符号化方式における最新のn個の音声フレームの音声符号を逆量子化して得られる逆量子化値を保持するバッファ、n個の逆量子化値を平均する平均値算出部、前記最初

の非音声フレームが検出されたとき、前記平均値を量子化する量子化器、を備え、量子化器の出力に基いて前記第2の非音声符号化方式における非音声符号を出力することを特徴とする付記15記載の音声符号変換装置。

【0093】(付記17) (1)第1の非音声符号化方式が、非音声区間における非音声信号の変化の度合が大きいフレームにおいてのみ非音声符号を伝送し、その他のフレームでは非音声符号を伝送せず、また、連続して非音声符号を伝送しない方式であり、(2)第2の非音声符号化方式が、非音声区間における所定フレーム数N毎に平均した非音声符号を伝送すると共に、その他のフレームでは非音声符号を伝送しない方式であり、更に、(3)第1の非音声符号化方式のフレーム長が、第2の非音声符号化方式のフレーム長の半分であるとき、前記非音声符号変換部は、第1の非音声符号化方式の連続する2×Nフレームにおける各非音声符号の逆量子化値を保持するバッファ、保持されている逆量子化値の平均値を演算する平均値算出部、平均値を量子化して第2の非音声符号化方式におけるNフレーム毎の非音声符号に変換する量子化器、Nフレーム毎以外のフレームについては、第1の非音声符号化方式の連続する2つのフレームの符号情報をフレームタイプに関係なく第2の非音声符号化方式の1つの非伝送フレームの符号情報に変換する手段、を備えたことを特徴とする付記14記載の音声符号変換装置。

【0094】(付記18) 音声区間から非音声区間に変化するとき、前記第2の非音声符号化方式が、変化点のフレームを含めて連続nフレームを音声フレームとみなして音声符号を伝送し、次のフレームは非音声符号を含まない最初の非音声フレームとしてフレームタイプ情報を伝送する場合、非音声符号変換部は、第1の非音声フレームの非音声符号を逆量子化して複数の要素符号の逆量子化値を発生する逆量子化器、予め定めた、あるいはランダムな複数の要素符号の逆量子化値を発生する手段、を備え、連続する2フレームの各要素符号の逆量子化値を第2音声符号化方式の量子化テーブルを用いてそれぞれ量子化して第2の音声符号化方式の1フレーム分の音声符号に変換して出力し、nフレーム分の第2音声符号化方式の音声符号を出力した後、非音声符号を含まない前記最初の非音声フレームのフレームタイプ情報を送出する、ことを特徴とする付記17記載の音声符号変換装置。

【0095】

【発明の効果】以上、本発明によれば、非音声符号化方法が異なる2つの音声通信システム間の通信において、送信側の非音声符号化方法で符号化した非音声符号(CN符号)をCN信号に復号しなくても受信側の非音声符号化方法に応じた非音声符号(CN符号)に変換することができ、高品質な非音声符号変換を実現できる。また、本発明によれば、送信側と受信側のフレーム長の相違やDTX制御の相違を考慮して非音声信号に復号することなく送信側の非音声符号(CN符号)を受信側の非音声符号(CN

符号)に変換することができ、高品質な非音声符号変換を実現できる。

【0096】また、本発明によれば音声フレームに加えて非音声圧縮機能によるSIDフレームおよび非伝送フレームに対しても正常な符号変換処理を行うことができる。これにより、従来の音声符号変換部で課題となっていた非音声圧縮機能を持つ音声符号化方式間での符号変換が可能となる。また、本発明によれば非音声圧縮機能の伝送効率向上効果を維持しつつ、さらに品質劣化と伝送遅延を抑えた異なる通信システム間の音声符号変換が可能となる。VoIPや携帯電話システムを始めとしてほとんどの音声通信システムでは非音声圧縮機能が用いられており、本発明の効果は大きい。

【図面の簡単な説明】

【図1】本発明の原理説明図である。

【図2】本発明の非音声符号変換の第1実施例の構成図である。

【図3】G.729AとAMRの処理フレームである。

【図4】AMRからG.729Aへのフレームタイプの変換制御手順である。

【図5】電力修正部の処理フローである。

【図6】本発明の第2実施例の構成図である。

【図7】本発明の第3実施例の構成図である。

【図8】音声区間での変換制御説明図である。

【図9】非音声区間での変換制御説明図である。

【図10】非音声区間での変換制御説明図(AMR8フレーム毎の変換制御)である。

【図11】本発明の第4実施例の構成図である。

【図12】第4実施例における音声符号変換部の構成図である。

【図13】音声→非音声変化点での変換制御説明図である。

【図14】非音声→音声変化点での変換制御説明図である。

【図15】従来技術1(タンデム接続)の説明図である。

【図16】従来技術2の説明図である。

【図17】従来技術2のより詳細な説明図である。

【図18】非音声圧縮機能の概念図である。

【図19】非音声圧縮機能の原理図である。

【図20】非音声圧縮機能の処理ブロック図である。

【図21】非音声圧縮機能の処理フローである。

【図22】非音声符号構成図である。

【図23】G.729AのDTX制御説明図である。

【図24】AMRのDTX制御(非ハングオーバ制御時)説明図である。

【図25】AMRのDTX制御(ハングオーバ制御時)説明図である。

【図26】従来技術において非音声圧縮機能を持つ場合の構成図である。

【符号の説明】

51a 符号化方式1の符号器

51b VAD部

52 フレームタイプ検出部

53 変換制御部

54 符号化方式2の復号器

60 非音声符号変換部

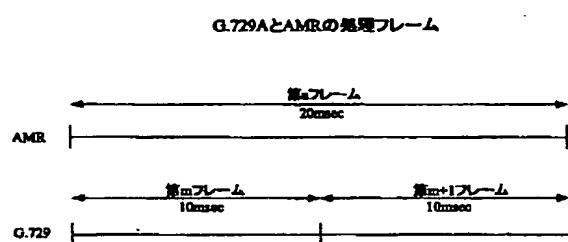
61 符号分離部

62₁~62n CN符号変換部

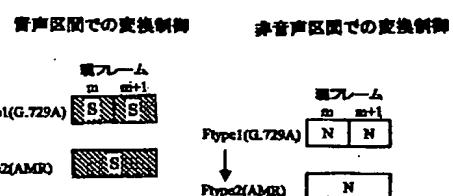
63 符号多重部

70 音声符号変換部

【図3】

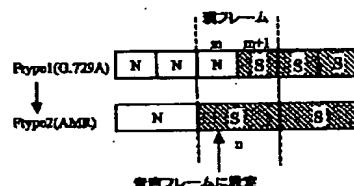


【図8】



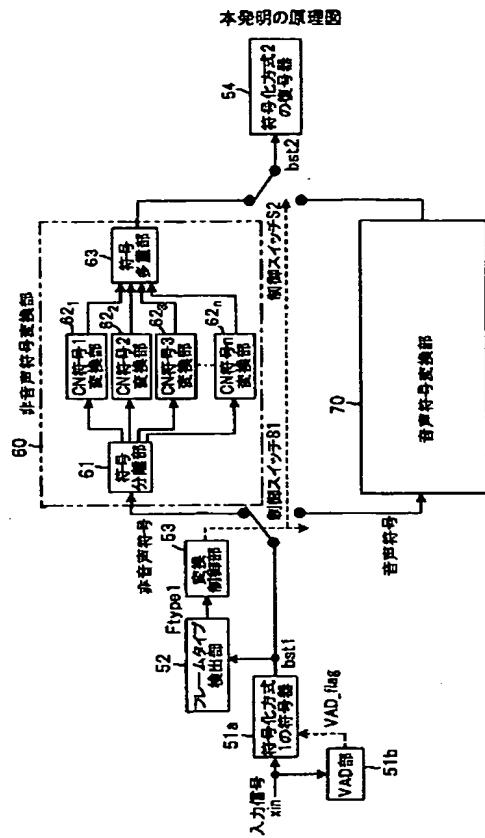
【図9】

非音声→音声変化点での変換制御

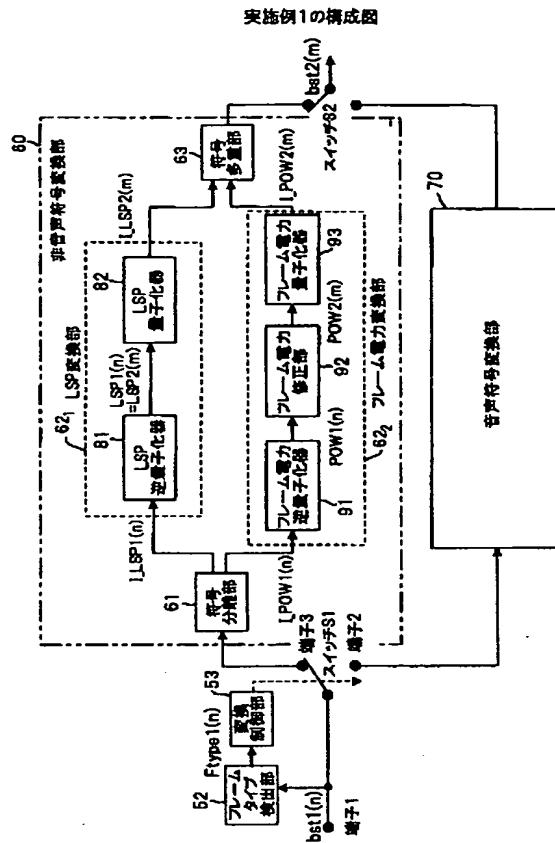


【図14】

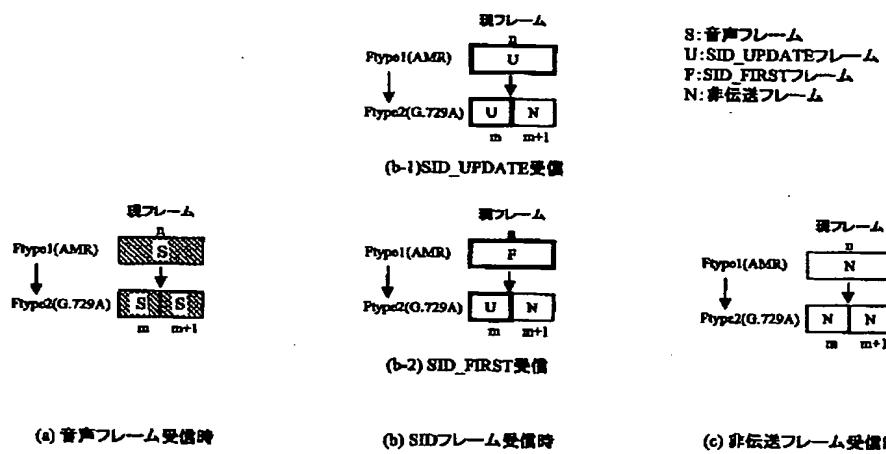
【図1】



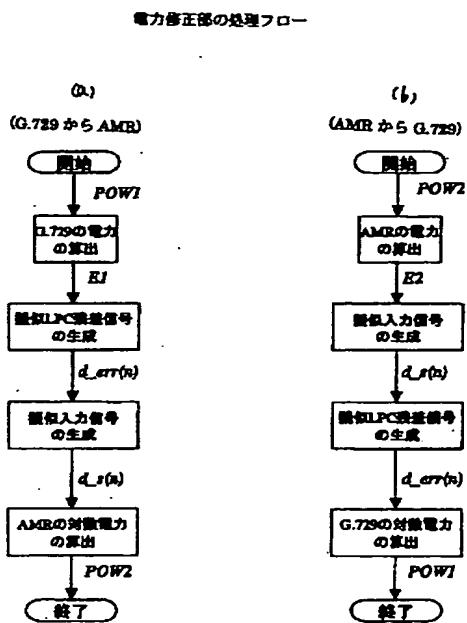
【図2】



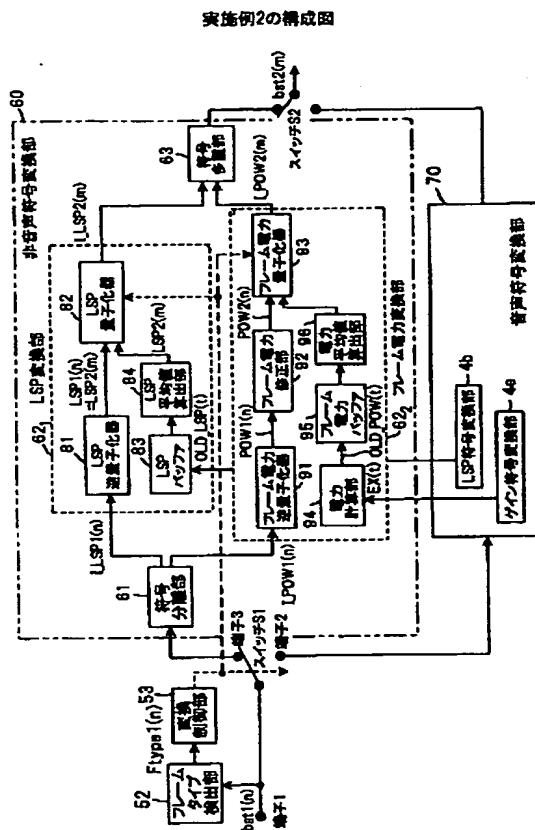
【図4】



【図5】

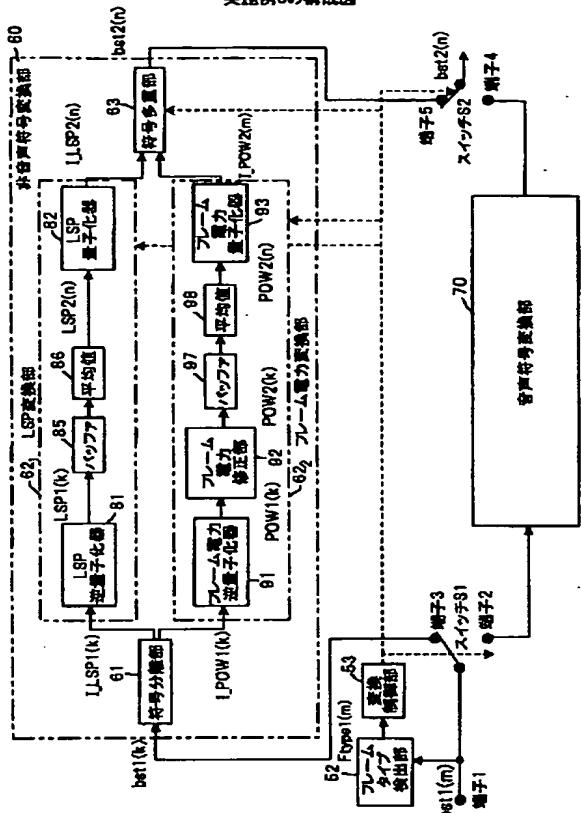


【図6】



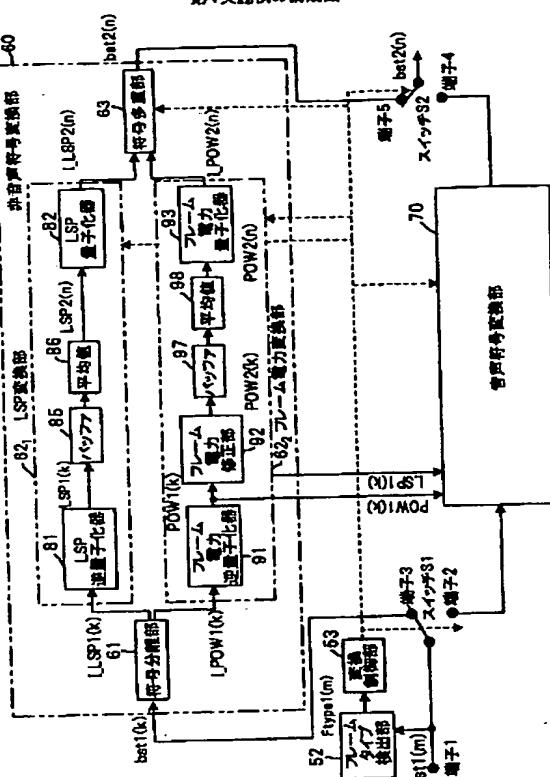
【図7】

実施例3の構成図

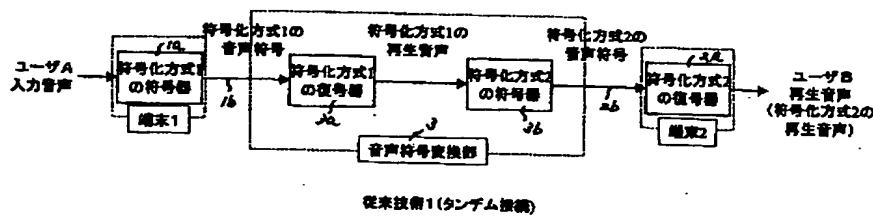


【図11】

第4実施例の構成図

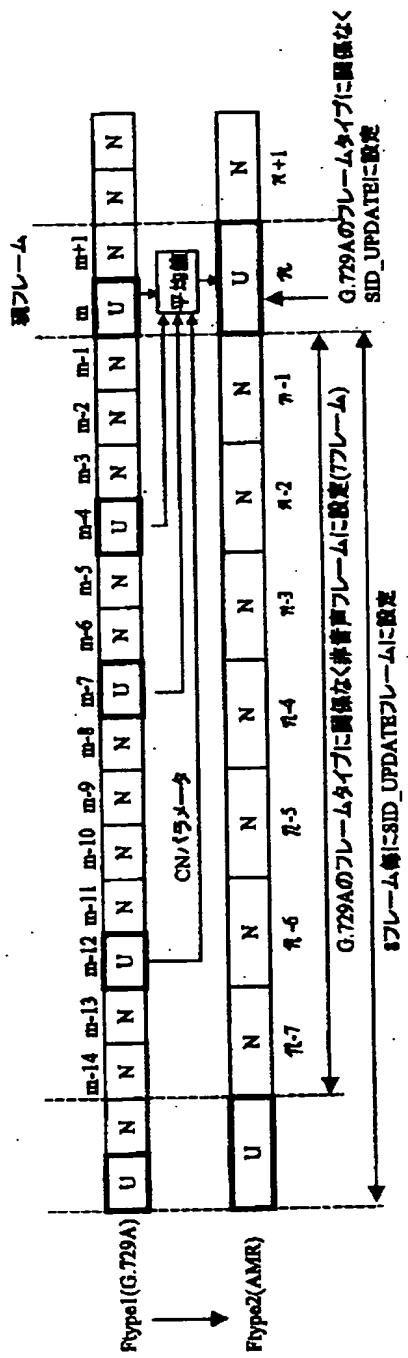


【図15】

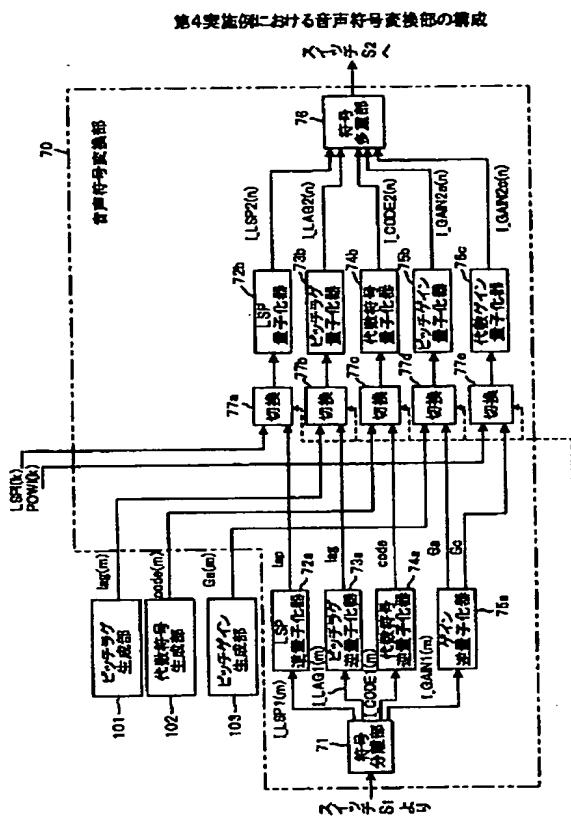


【図10】

非音区間での変換制御(AMRの8フレーム毎の変換制御)



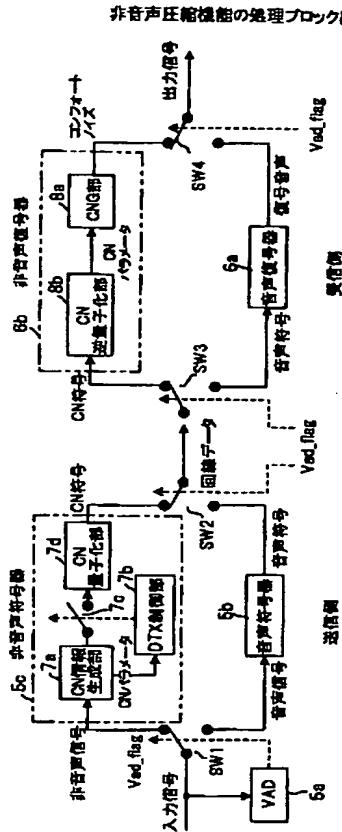
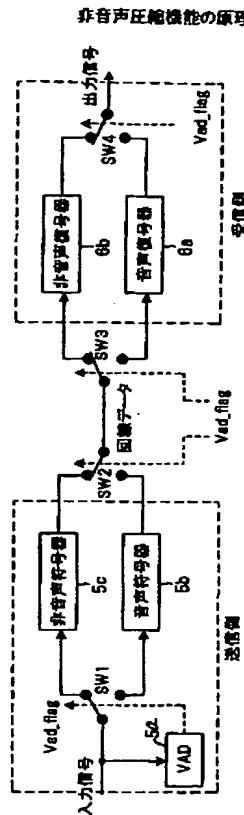
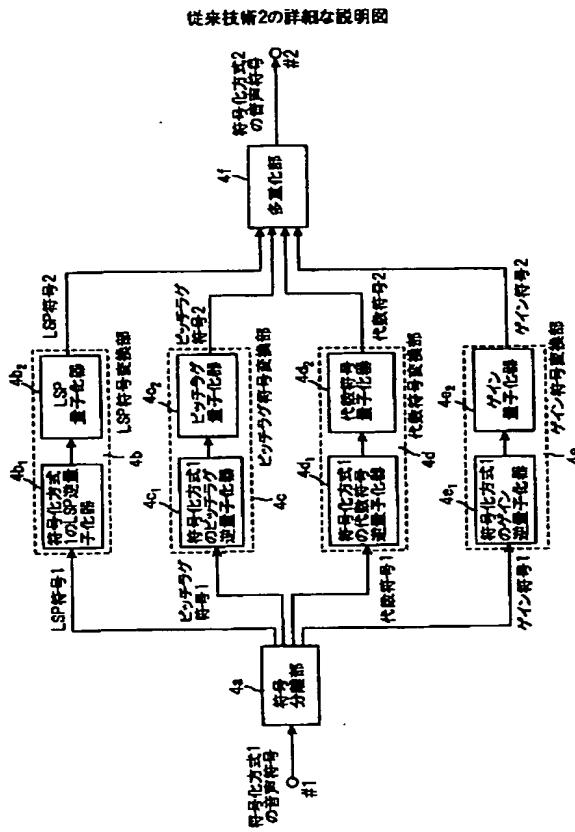
【図12】



〔図17〕

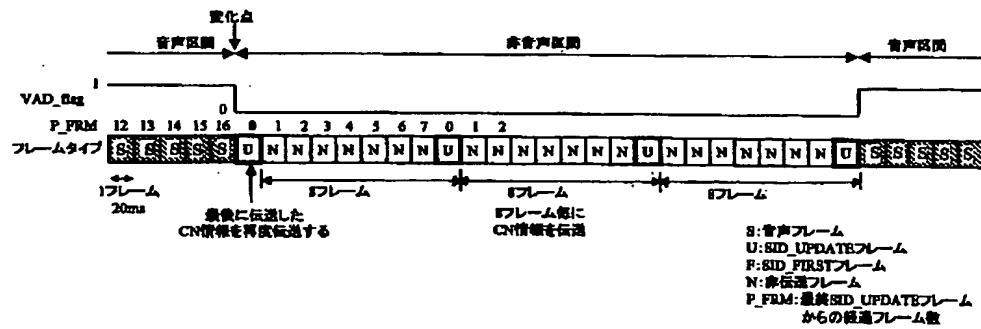
〔図19〕

〔図20〕

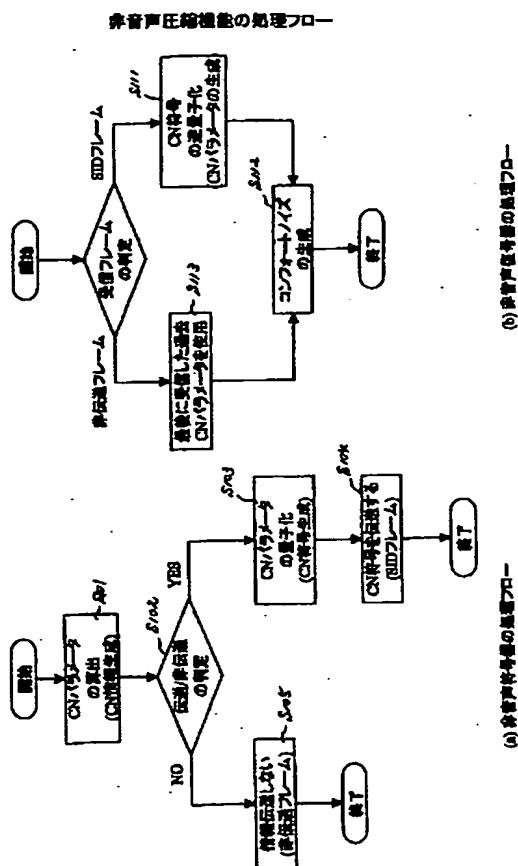


【图24】

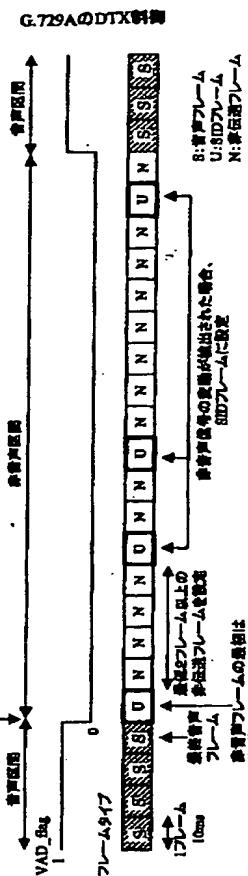
AMRのDTX制御(非ハンギングオーバー制御時)



【图21】

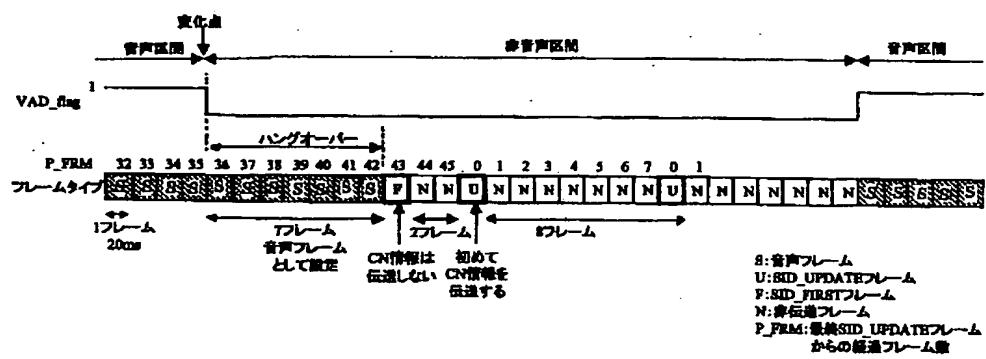


〔図23〕

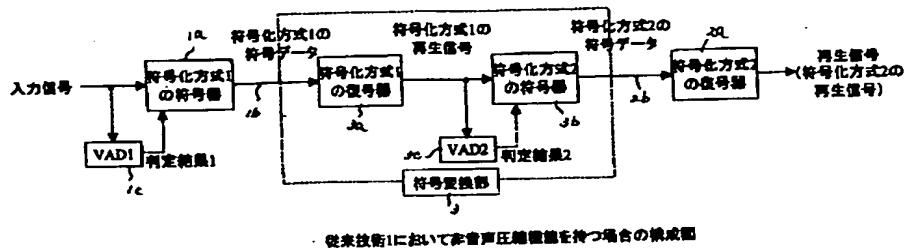


【图25】

AMRのDTX制御（ハングオーバー制御時）



【図26】



従来技術において高音声圧縮機能を持つ場合の構成図

フロントページの続き

(72)発明者 大田 恒士
神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内
(72)発明者 鈴木 政直
神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

F ターム(参考) 5D045 CC10 DA20
5J064 AA01 BA01 BB03 BC01 BC16
BC21 BC25 BC26 BD02
5K028 AA01 AA14 BB04 CC05 KK23
LL29 MM08 SS04 SS05 SS14